arXiv:1203.3182v1 [math.OC] 14 Mar 2012

Maximum Principle for General Controlled Systems Driven by Fractional Brownian Motions

Yuecai Han*, Yaozhong Hu† and Jian Song

Abstract

We obtain a maximum principle for stochastic control problem of general controlled stochastic differential systems driven by fractional Brownian motions (of Hurst parameter $H > 1/2$). This maximum principle specifies a system of equations that the optimal control must satisfy (necessary condition for the optimal control). This system of equations consists of a backward stochastic differential equation driven by both fractional Brownian motion and the corresponding underlying standard Brownian motion. In addition to this backward equation, the maximum principle also involves the Malliavin derivatives. Our approach is to use conditioning and Malliavin calculus. To arrive at our maximum principle we need to develop some new results of stochastic analysis of the controlled systems driven by fractional Brownian motions via fractional calculus. Our approach of conditioning and Malliavin calculus is also applied to classical system driven by standard Brownian motion while the controller has only partial information. As a straightforward consequence, the classical maximum principle is also deduced in this more natural and simpler way.

1 Introduction

Fix a finite time horizon $T \in (0, \infty)$. Let $(\Omega, \mathcal{F}, P)$ be a basic probability space equipped with a right continuous filtration $(\mathcal{F}_t)_{0\le t\le T}$ satisfying the usual conditions ([9]). Let $B^H = (B_1^H(t), \cdots, B_m^H(t),\ 0 \le t \le T)$ be an m-dimensional fractional Brownian motion of Hurst parameter $H \in [\frac{1}{2}, 1)$ (It is straightforward except notational complexity to allow H to be different for different fractional Brownian motions). This means that $B_j^H(t)$, $j = 1, 2, \cdots, m$, are independent, continuous, mean 0 Gaussian processes with the following covariance

$$\mathbb{E}\left(B_i^H(t)B_j^H(s)\right) = \frac{1}{2}\delta_{ij}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \tag{1.1}$$

*Y. Han is partially supported by NSFC grant 11071101, NSF of Jilin grant 20101594 and the Science Research Fund grant 200903281 of Jilin University

†Y. Hu is partially supported by a grant from the Simons Foundation #209206.

Keywords: Stochastic optimal control, backward differential equations, maximum principle, fractional Brownian Motion, controlled stochastic differential systems driven fractional Brownian Motion, Malliavan calculus, partial information stochastic control.

where $\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$ is the Kronecker symbol.

This process has been applied in many fields such as hydrology, climatology, economics, internet traffic analysis, finance, and many other fields. The stochastic analysis associated with fractional Brownian motions has been extensively studied recently. The stochastic differential equations driven by fractional Brownian motions have also been considered by many researchers through several approaches, see for example, through general rough path analysis [7], [12], [16], [25] or through fractional calculus [17], [18], [28]. In particular, we refer to the references therein.

Since stochastic control is a main tool of applications of stochastic analysis it is natural to consider the problem of stochastic control of systems driven by fractional Brownian motions. Along this direction there have been already some work. In [15], [21] (see also [1]) some specific stochastic control problems relevant to mathematical finance have been investigated. There is also a general sufficient condition of optimal control for general control problems in [20]. The explicit optimal linear Markov control was obtained in [22] by using the technique of completing squares and by using the Riccati equations.

However, the problem of optimal control for general stochastic systems driven by fractional Brownian motions is far away from being considered as resolved. In fact, there has been a lack of necessary conditions in a more general setting. In this paper, we shall fill this gap. More precisely, we shall obtain a set of necessary conditions that the optimal control must satisfy.

The theory of stochastic control of systems driven by standard Brownian motions is very rich and has found many applications. There have been mainly two general approaches toward the solutions. One is the Hamilton-Jacobi-Bellman dynamic programming which results in a highly nonlinear Hamilton-Jacobi-Bellman equation. The study of viscosity solutions of this type of equations has experienced an explosive growth in recent years. See [11], [35], and the references therein for the stochastic control related development. The study of viscosity solutions of the nonlinear Hamilton-Jacobi-Bellman equation has become one main stream of partial differential equations. see [8], [24], and many more other references for general discussion. Another approach in stochastic control is the Pontryagin's maximum principle. Starting with [4], [5], and [6], backward stochastic differential equations (abbreviated as BSDEs) has been used to describe the necessary (and sufficient) conditions that the optimal control must satisfy. We also refer to [11], [29], [35] and the references therein for some other work.

The approach of Bellman dynamic programming heavily depends on the semigroup property of the underlying system (namely, the Markov property of the underlying controlled stochastic processes). In fact, one can also obtain the Hamilton-Jacobi-Bellman equation for more general Markov processes, see for example, [26]. However, the fractional Brownian motions are not Markov except in the case of Brownian motion ($H = 1/2$). Thus, it is natural to concentrate on extending the Pontryagin's maximum principle to controlled system driven by

fractional Brownian motions. The first main task is to find the appropriate backward stochastic differential equations (an extension of the Riccati equation). In this work, we obtain the backward stochastic differential equations in a natural way, through the idea of conditioning. This type of backward stochastic differential equations involved terms driven both by fractional Brownian motion and by the underlying standard Brownian motion. In the classical standard Brownian motion case, researchers usually obtain this backward stochastic differential equation by the duality approach for which one has to know the form of the BSDE in advance. Our approach is motivated by a recent work [19] on linear BSDE driven by standard Brownian motion by using Malliavin calculus. We are excited about this approach since it is very natural: the BSDE is deduced naturally without prior knowledge of the form of the BSDE!

Our approach is also new in the classical setting of the controlled systems driven by standard Brownian motions. The advantage of our approach in the classical setting of standard Brownian motion is that it also works for stochastic control with *partial information*. Thus, we also present our approach to deduce the maximum principle in classical case, first with partial information and then to give an alternative way to deduce the classical maximum principle with complete information. This is done in Section 3.

To deduce the maximum principle for the controlled stochastic system driven by fractional Brownian motions, we need more results on stochastic analysis of the controlled systems, which has not been studied yet. In particular, we need to have the uniform Hölder continuity of the solutions, and the differentiability of the solution with respect to the control. These results are of interest themselves. We shall present these new results in Section 4. In Section 5.1, using the idea in Section 3, we obtain a necessary condition that the optimal control must satisfy when the controller has only partial information. The stochastic control problem with partial information is very important in finance, while not much theory has been developed yet. However, we refer to the work [20] (for partial information linear quadratic control) and the references therein. Section 5.2 aims to simplify the condition obtained in Section 5.1 when the controller has complete information available. The condition leads naturally to a new type of backward stochastic differential equation driven by the underlying standard Brownian motion and by the fractional Brownian motion. Besides the complexity of the BSDE which involves fractional Brownian motion and the underlying standard Brownian motion, the system of equations of maximum principle also involves the Malliavin derivatives. This system of equations of maximum principle is very complex. However, this is expected since the problem is much more complicated now. In the case of controlled system driven by standard Brownian motion, the theory obviously reduces to the classical maximum principle as demonstrated in Section 3.

To obtain our maximum principle we need some additional results on fractional calculus and Malliavin calculus. In Section 2, we introduce some notations and obtain some new results that we shall use. We also recall some necessary notations from [1], [10], [14], [16], [17], and [18] to establish some new results for controlled system driven by fractional Brownian motions. In particular we

shall establish some new important identity which is necessary in our approach.

It is interesting to have some examples that can be solved by using our new maximum principle. However, as it is well-known, the explicit solutions of stochastic control are always difficult to obtain even in the classical Brownian motion case. See however, [3] for some general discussion. It is expected that the problem will be more complex. In addition, this paper is already long. So, we shall discuss some particular control problems with this approach in the future.

2 Fractional calculus and Malliavin calculus

2.1 Fractional calculus

In this section we recall some results from fractional calculus. Let $a, b \in \mathbb{R}$ with $a < b$ and let $\alpha > 0$. The left-sided (and right-sided) fractional Riemann-Liouville integrals of integrable function f is defined by

$$I_{a+}^{\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_a^t (t-s)^{\alpha-1} f(s)\, ds\,, \quad a \le t \le b\,,$$

[and

$$I_{b-}^{\alpha} f(t) = \frac{(-1)^{-\alpha}}{\Gamma(\alpha)} \int_t^b (s-t)^{\alpha-1} f(s)\, ds,$$

respectively], where $(-1)^{-\alpha} = e^{-i\pi\alpha}$ and $\Gamma(\alpha) = \int_0^\infty r^{\alpha-1} e^{-r} dr$ is the Euler gamma function. The Weyl derivatives are defined as

$$D_{a+}^{\alpha} f(t) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(t)}{(t-a)^{\alpha}} + \alpha \int_a^t \frac{f(t) - f(s)}{(t-s)^{\alpha+1}} ds \right) \tag{2.1}$$

and

$$D_{b-}^{\alpha} f(t) = \frac{(-1)^{\alpha}}{\Gamma(1-\alpha)} \left(\frac{f(t)}{(b-t)^{\alpha}} + \alpha \int_t^b \frac{f(t) - f(s)}{(s-t)^{\alpha+1}} ds \right) \tag{2.2}$$

where $a \le t \le b$.

For any $\beta \in (0,1)$, we denote by $C^{\beta}(a,b)$ the space of β-Hölder continuous functions on the interval $[a,b]$. We will make use of the notation

$$\|x\|_{a,b,\beta} = \sup_{a \le \theta < r \le b} \frac{|x_r - x_\theta|}{|r-\theta|^{\beta}},$$

and

$$\|x\|_{a,b,\infty} = \sup_{a \le r \le b} |x_r|\,.$$

When a and b are clear, then we shall use $\|x\|_{\beta} = \|x\|_{a,b,\beta}$ and $\|x\|_{\infty} = \|x\|_{a,b,\infty}$.

It is clear that when $f \in C^{\beta}(a,b)$ with $\beta > \alpha$, then both $D_{a+}^{\alpha} f(t)$ and $D_{b-}^{\alpha} f(t)$ exist and we have

$$\begin{cases} \left| D_{a+}^{\alpha} f(t) \right| \le C \|f\|_{\beta} |t-a|^{\beta-\alpha} & \text{if } f(a) = 0 \\ \\ \left| D_{b-}^{\alpha} f(t) \right| \le C \|f\|_{\beta} |b-t|^{\beta-\alpha} & \text{if } f(b) = 0 \,. \end{cases} \tag{2.3}$$

The following fractional integration by parts formula and its consequence are needed in the sequel (see [34] for a proof).

Proposition 2.1 *Suppose that $f \in C^{\lambda}(a,b)$ and $g \in C^{\mu}(a,b)$ with $\lambda + \mu > 1$. Let $\lambda > \alpha$ and $\mu > 1 - \alpha$. Then the Riemann Stieltjes integral $\int_a^b f dg$ exists and it can be expressed as*

$$\int_a^b f dg = (-1)^{\alpha} \int_a^b D_{a+}^{\alpha} f(t)\, D_{b-}^{1-\alpha} g_{b-}(t)\, dt \,, \tag{2.4}$$

where $g_{b-}(t) = g(t) - g(b)$.

From the definition (2.1) of $D_{a+}^{\alpha} f(t)$ and (2.3), we have immediately

Proposition 2.2 *Suppose that $f \in C^{\lambda}(a,b)$ and $g \in C^{\mu}(a,b)$ with $\lambda + \mu > 1$. Let $\lambda > \alpha$ and $\mu > 1 - \alpha$. Then the Riemann Stieltjes integral $\int_a^b f dg$ exists and*

$$\begin{aligned} \left| \int_a^b f dg \right| \quad \le \quad & C\|g\|_{\mu} \int_a^b \frac{|f(r)|}{(r-a)^{\alpha}} (b-r)^{\alpha+\mu-1} dr \\ & + C\|g\|_{\mu} \int_a^b \int_a^r \frac{|f(r) - f(\tau)|}{|r-\tau|^{\alpha+1}} (b-r)^{\alpha+\mu-1} d\tau dr \,. \end{aligned} \tag{2.5}$$

2.2 Stochastic calculus for fractional Brownian motions

Let $(W(t) = (W_1(t), \cdots, W_m(t)), 0 \le t \le T)$ be an m-dimensional standard Brownian motion. Let

$$Z_H(t,s) = \kappa_H \left[\left(\frac{t}{s}\right)^{H-\frac{1}{2}} (t-s)^{H-\frac{1}{2}} - (H - \frac{1}{2}) s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u-s)^{H-\frac{1}{2}} du \right] \tag{2.6}$$

with

$$\kappa_H = \sqrt{\frac{2H\Gamma(\frac{3}{2} - H)}{\Gamma(H + \frac{1}{2})\Gamma(2 - 2H)}} \,. \tag{2.7}$$

and define

$$B_j^H(t) = \int_0^t Z_H(t,s) dW_j(s) \,,\ 0 \le t < \infty \,. \tag{2.8}$$

Then from (2.2)-(2.4) of [14] we see that $B^H(t) = (B_1^H(t), \cdots, B_m^H(t)), 0 \le t \le T)$ is an m dimensional fractional Brownian motion. This means that

$B_j^H(t)\,, j=1,\cdots,m$ are independent Gaussian processes with mean 0 and variance given by

$$\mathbb{E}\left(B_j^H(t)B_i^H(s)\right)=\frac{1}{2}\delta_{ij}\left(r^{2H}+s^{2H}-|t-s|^{2H}\right)\,.$$

The stochastic integral with respect to B_j^H can be defined in a similar way as in [14]. We shall use the results in [14].

Recall the operators $\mathbb{\Gamma}_{H,T}^*$ and $\mathbb{B}_{H,T}^*$ (see the equations (5.21) and (5.35) of [14])

$$\mathbb{\Gamma}_{H,T}^*f(t):=(H-\frac{1}{2})\kappa_H t^{\frac{1}{2}-H}\int_t^T u^{H-\frac{1}{2}}(u-t)^{H-\frac{3}{2}}f(u)du\,,\quad 0\le t\le T\,. \tag{2.9}$$

and

$$\mathbb{B}_{H,T}^*f(t)=-\frac{2H\kappa_1}{\kappa_H}t^{\frac{1}{2}-H}\frac{d}{dt}\int_t^T(u-t)^{\frac{1}{2}-H}u^{H-\frac{1}{2}}f(u)du\,, \tag{2.10}$$

where κ_H is defined in (2.7) and

$$\kappa_1=\frac{1}{2H\Gamma(H-\frac{1}{2})\Gamma(\frac{3}{2}-H)}.$$

Let $\xi_1,\,\cdots,\,\xi_k,\,\cdots$ be an ONB of $L^2([0,T])$ such that $\xi_k,\ k=1,2,\cdots$ are smooth functions on $[0,T]$. We denote $\tilde{\xi}_{j,l}=\int_0^T\xi_j(t)dW_l(t)\,,$ where $j=1,2,\cdots,$ and $l=1,\cdots,m$. Let $\mathcal{P}$ be the set of all polynomials of the standard Brownian motion W over interval $[0,T]$. Namely, $\mathcal{P}$ contains all elements of the form

$$F(\omega)=f\left(\tilde{\xi}_{j_1,l_1}\,,\cdots,\tilde{\xi}_{j_n,l_n}\right)\,, \tag{2.11}$$

where f is a polynomial of n variables. If F is of the above form (2.11), then its Malliavin derivative D_sF is defined as

$$D_s^lF=\sum_{i=1}^n\frac{\partial f}{\partial x_i}\left(\tilde{\xi}_{j_1,l_1}\,,\cdots,\tilde{\xi}_{j_n,l_n}\right)\xi_{j_i}(s)I_{\{l_i=l\}}\,,\quad 0\le s\le T\,. \tag{2.12}$$

For any $F\in\mathcal{P}$, we denote the following norm

$$\|F\|_{1,p}:=\|F\|_p+\sum_{l=1}^m\left[\mathbb{E}\left(\int_{[0T]}|D_t^lF|^2dt\right)^{p/2}\right]^{1/p}\,.$$

It is easy to check that $\|\cdot\|_{1,p}$ is a norm $p\in(1,\infty)$. Let $\mathbb{D}_{1,p}$ denote the Banach space obtained by completing $\mathcal{P}$ under the norm $\|\cdot\|_{1,p}$. We can certainly define the higher derivatives. But we only need the first derivatives in this paper.

For the above ξ_j's we can define $\eta_j=\mathbb{B}_{H,T}^*\xi_j$ and we denote by $\mathcal{P}^H$ the set of all polynomial functionals of $\tilde{\eta}_{j,l}:=\int_0^T\eta_j(t)dB_l^H(t)$. For an element G in $\mathcal{P}^H$ of the following form

$$G(\omega)=g\left(\tilde{\eta}_{j_1,l_1}\,,\cdots,\tilde{\eta}_{j_n,l_n}\right)\,, \tag{2.13}$$

where g is a polynomial of n variables we define its Malliavin derivative $D_s^{H,l}G$ by

$$D_s^{H,l}G = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}\left(\tilde{\eta}_{j_1,l_1}, \cdots, \tilde{\eta}_{j_n,l_n}\right) \eta_{j_i}(s) I_{\{l_i=l\}}, \quad 0 \le s \le T. \tag{2.14}$$

Similarly, we can define $\|\cdot\|_{H,1,p}$ and $\mathbb{D}_{H,1,p}$.

For fractional Brownian motions, another different Malliavin derivative is also very useful.

$$\mathbb{D}_s^l G = \int_0^T \phi(s-r) D_r^{H,l} G dr, \tag{2.15}$$

where

$$\phi(s) = H(2H-1)|s|^{2H-2}, \quad 0 \le s \le T. \tag{2.16}$$

It is well-known from ([14], Theorem 6.23) that

Proposition 2.3 *Let $\psi : [0,T] \otimes (\Omega, \mathcal{F}, P^H) \to \mathbb{R}$ be jointly measurable and let $F \in D_{H,1,2}$. Then*

$$\mathbb{E}\left\{F \int_a^b \psi_t dB_j^H(t)\right\} = \int_a^b \mathbb{E}\left(\left[\mathbb{D}_t^j F\right] \psi_t\right) dt \tag{2.17}$$

The following proposition from ([10], Theorem 3.9) will also be used in this the sequel.

Proposition 2.4 *Let ψ be a jointly measurable stochastic process which has Malliavin derivative and let the following hold:*

$$\mathbb{E}\left[\int_a^b \int_a^b |\psi_s \psi_t| \phi(s-t) ds dt + \int_a^b \int_a^b |\mathbb{D}_s \psi_t|^2 ds dt\right] < \infty.$$

Then

$$\int_a^b \psi(t) d^\circ B_j^H(t) = \int_a^b \psi(t) dB_j^H(t) + \int_a^b \mathbb{D}_t^j \psi(t) dt. \tag{2.18}$$

The following proposition is easy consequence of the above identity (2.17) and the relation between path-wise integral and the stochastic integral obtained by using the Wick product (2.18).

Proposition 2.5 *If F satisfies the condition in Proposition 2.3 and ψ satisfies the conditions in Propositions 2.3 and 2.4, then*

$$\mathbb{E}\left\{F \int_a^b \psi_t d^\circ B_j^H(t)\right\} = \int_a^b \mathbb{E}\left(\mathbb{D}_t^j \left[F \psi_t\right]\right) dt. \tag{2.19}$$

If F is a nice functional of the fractional Brownian motions $B^H=(B_1^H,\cdots,B_m^H)$, then it is also a functional of standard Brownian motions $W=(W_1,\cdots,W_m)$. So for this F we can compute both the Malliavin derivatives D_s^l and $D_s^{H,l}$. The relation between these two Malliavin derivatives are useful later in this paper. Next, we shall find this relation.

If f is a nice function on $[0,T]$, then $\int_0^T f(t)dW_j(t)$ and $\int_0^T f(t)dB_j^H(t)$ are well-defined and from page 45 of [14], we have

$$\int_0^T f(t)dB_j^H(t)=\int_0^T (\mathbb{\Gamma}_{H,T}^* f)(t)dW_j(t)$$

and

$$\int_0^T f(t)dW_j(t)=\int_0^T (\mathbb{B}_{H,T}^* f)(t)dB_j^H(t)\,,$$

where $\mathbb{\Gamma}_{H,T}^*$ and $\mathbb{B}_{H,T}^*$ are defined by (2.9) and (2.10). If F is given by (2.11), then F can also be represented by

$$F=f\left(\int_0^T (\mathbb{B}_{H,T}^*\xi_{j_1})(t)dB_{l_1}^H(t)\,,\cdots,\int_0^T (\mathbb{B}_{H,T}^*\xi_{j_n})(t)dB_{l_1}^H(t)\right)\,.$$

Thus as a functional of B^H, its Malliavin derivative D_sF is defined as

$$\begin{aligned} D_s^{H,l}F &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}\left(\int_0^T (\mathbb{B}_{H,T}^*\xi_{j_1})(t)dB_{l_1}^H(t)\,,\cdots,\int_0^T (\mathbb{B}_{H,T}^*\xi_{j_n})(t)dB_{l_1}^H(t)\right)\\ &\qquad (\mathbb{B}_{H,T}^*\xi_{j_i})(s)I_{\{l_i=l\}}\,,\quad 0\le s\le T\,. \end{aligned}$$

Thus for $F\in\mathcal{P}$, we have

$$\begin{aligned} D_s^{H,l}F &= \mathbb{B}_{H,T}^* D_\cdot^l F(s)\\ &= -\frac{2H\kappa_1}{\kappa_H}s^{\frac12-H}\frac{d}{ds}\int_s^T (u-s)^{\frac12-H}u^{H-\frac12}D_u^lFdu \end{aligned}$$

and

$$\mathbb{D}_t^lF = -\frac{2H\kappa_1}{\kappa_H}H(2H-1)\int_0^T s^{\frac12-H}|t-s|^{2H-2}\left(\frac{d}{ds}\int_s^T (u-s)^{\frac12-H}u^{H-\frac12}D_u^lFdu\right)ds\,.$$

By a limiting argument, we have the following

Proposition 2.6 *If $F\in\mathbb{D}_{1,p}\cap\mathbb{D}_{H,1,p}$, then*

$$D_s^{H,l}F = -\frac{2H\kappa_1}{\kappa_H}s^{\frac12-H}\frac{d}{ds}\int_s^T (u-s)^{\frac12-H}u^{H-\frac12}D_u^lFdu \tag{2.20}$$

$$\mathbb{D}_t^lF = c_{1,H}\int_0^T s^{\frac12-H}|t-s|^{2H-2}\left(\frac{d}{ds}\int_s^T (u-s)^{\frac12-H}u^{H-\frac12}D_u^lFdu\right)ds\,, \tag{2.21}$$

where $c_{1,H}=-\frac{2H^2(2H-1)\kappa_1}{\kappa_H}$.

Proposition 2.7 *Let $F_t = \sum_{j=1}^m \int_0^t f_j(s)dB_j^H(s)$ and $G_t = \sum_{j=1}^m \int_0^t g_j(s)dW_j(s)$, where $f_1, \cdots, f_m$ satisfy the conditions in Proposition 2.4 and let $g_1, \cdots, g_m$ be continuous adapted processes. Then*

$$d(F_tG_t) = FdG_t + G_tdF_t\,. \tag{2.22}$$

Proof We can prove this proposition in the same way as the proof for Theorem 2.1 of [22]. ■

3 Systems driven by Brownian motions

In this section, we shall use our approach of conditioning and Malliavin calculus to deduce the maximum principle for partially observed controlled system driven by standard Brownian motions. The complete information case will also be deduced. In this classical case, our approach seems to be new, straightforward and simpler.

3.1 General stochastic control with partial information

In this subsection we shall obtain a maximum principle when only partial information is available. This means that the control u_t may not necessarily depends on full information determined by the Brownian motions $\{W(t) : 0 \le t \le T\}$. We assume in this subsection that $(\mathcal{G}_t)_{0\le t\le T}$ is any right continuous filtration which is contained in the σ-algebra $(\mathcal{F}_t)_{0\le t\le T}$ generated by $\{W(t) : 0 \le t \le T\}$.

The space of admissible controls is defined as

$$\begin{aligned} U[0,T] \quad &\triangleq \quad \Big\{u : [0,T]\times\Omega \to \mathbb{R}^d \;|\;\; u \text{ is } \mathcal{G}\text{-adapted stochastic process} \\ &\qquad\qquad \text{and} \quad \mathbb{E}\left(\int_0^T |u(t)|^2 dt\right) < +\infty\Big\}. \end{aligned}$$

To describe our stochastic control problem, we need to introduce four functions b, σ, l, and h. They are assumed to satisfy the following conditions.

(H1) The functions $b : [0,T]\times\mathbb{R}^n\times\mathbb{R}^d \to \mathbb{R}^n$, $\sigma = (\sigma^1, \cdots, \sigma^m) : [0,T]\times\mathbb{R}^n\times\mathbb{R}^d \to \mathbb{R}^{n\times m}$, and $h : \mathbb{R}^n \to \mathbb{R}$ are continuous with respect variables t, x, and u and are continuously differentiable with respect to x, u for all $t \in [0,T]$.

Denote

$$b_x(t,x,u) = \left(\frac{\partial b_i(t,x,u)}{\partial x_j}\right)_{1\le i,j\le n}, \quad b_u(t,x,u) = \left(\frac{\partial b_i(t,x,u)}{\partial u_j}\right)_{1\le i\le n, 1\le j\le d}$$

(H2) We assume that there is a constant $C > 0$ such that

$$\sup_{t\in[0,T],x\in\mathbb{R}^n,u\in\mathbb{R}^d} |b_x(t,x,u)|+|b_u(t,x,u)|+\sum_{j=1}^m |\sigma_x^j(t,x,u)|+\sum_{j=1}^m |\sigma_u^j(t,x,u)| \le C\,.$$

(H3) We assume that $l : [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable with bounded derivatives.

[In what follows throughout this paper, we shall use C to denote a generic constant which may have different value at different occurrences.]

The controlled stochastic control system is described as the following stochastic differential equation:

$$\left\{ \begin{array}{rcl} dx(t) & = & b(t,x(t),u(t))dt + \sum_{j=1}^{m} \sigma^j(t,x(t),u(t))dW_j(t), \\ x(0) & = & x_0\,, \end{array} \right. \tag{3.1}$$

where x_0 is a given vector in $\mathbb{R}^n$. For a given $u \in U[0,T]$ the existence and uniqueness of the solution $x^u(t)$ to the above equation follows from standard theory of stochastic differential equations. For simplicity, we sometimes omit the explicit dependence of $x(t)$ on u, namely, we write $x(t) = x^u(t)$. The cost functional we shall deal with is

$$J(u(\cdot)) = J(x^u(\cdot), u(\cdot)) = \mathbb{E}\left\{ \int_0^T l(t, x^u(t), u(t))dt + h(x^u(T)) \right\}. \tag{3.2}$$

The first optimal control problem studied in this paper is to minimize the cost functional $J(u(\cdot))$ over $U[0,T]$. This means that we want to find the optimal control $u^*(\cdot) \in U[0,T]$ satisfying

$$J(u^*(\cdot)) = \inf_{u(\cdot)\in U[0,T]} J(u(\cdot))\,. \tag{3.3}$$

If such an optimal control $u^*(\cdot)$ exists, then the corresponding state $x^*(\cdot) = x^{u^*}(\cdot)$ is called an optimal state process. $(x^*(\cdot), u^*(\cdot))$ is called an optimal pair for the optimal control problem described by (3.1), (3.2) and (3.3).

Assume $(x^*(\cdot), u^*(\cdot))$ is an optimal pair. Namely, $J(u(\cdot))$ attains its minimum at $u^*(\cdot)$. From the general theory of functional analysis, we know that $u^*(\cdot)$ is a critical point of $J(u(\cdot))$. We are interested in finding a necessary condition satisfied by $(x^*(\cdot), u^*(\cdot))$. For any $\varepsilon \in \mathbb{R}$ and $v(\cdot) \in U[0,T]$, let $u^\varepsilon(\cdot) = u^* + \varepsilon v(\cdot)$. It is easy to see that $u^\varepsilon(\cdot) \in U[0,T]$, and there is a unique $x^\varepsilon(\cdot)$ satisfying the state equations (3.1) with control u replaced by $u^\varepsilon(\cdot)$.

Denote

$$\begin{array}{rcl} b_x^*(t) & = & b_x(t, x^*(t), u^*(t))\,, \quad b_u^*(t) = b_u(t, x^*(t), u^*(t))\,, \\ \sigma_x^{j,*}(t) & = & \sigma_x^{j,*}(t, x^*(t), u^*(t))\,, \quad \sigma_u^{j,*}(t) = \sigma_u^{j,*}(t, x^*(t), u^*(t))\,. \end{array} \tag{3.4}$$

The following lemma can be proved easily and it is necessary in obtaining the maximum principle for the stochastic control problem (3.1)-(3.3).

Lemma 3.1 *The limit* $y(t) = \lim_{\varepsilon\to 0} \frac{x^\varepsilon(t) - x^*(t)}{\varepsilon}$ *exists in* L^2 *and* $y(t)$ *satisfies the following equations:*

$$\left\{ \begin{array}{rcl} dy(t) & = & [b_x^*(t)y(t) + b_u^*(t)v(t)]dt + \sum_{j=1}^{m} [\sigma_x^{j,*}(t)y(t) + \sigma_u^{j,*}(t)v(t)]dW_j(t), \\ y(0) & = & 0. \end{array} \right. \tag{3.5}$$

We need to solve the above linear stochastic differential equations (with random coefficients). In particular, we need to express y in an explicit form of v. For this reason we consider the following linear matrix-valued stochastic differential equations.

$$\begin{cases} d\Phi(t) &= b_x^*(t)\Phi(t)dt + \sum_{j=1}^m \sigma_x^{j,*}(t)\Phi(t)dW_j(t), \\ \Phi(0) &= I, \end{cases} \tag{3.6}$$

From the basic theory of stochastic differential equations, it is well-known that this equation has a unique solution, denoted by $\Phi(t)$. It is easy to verify that $\Phi^{-1}(t)$ exists and is the unique solution of the following stochastic differential equations ([35]):

$$\begin{cases} d\Phi^{-1}(t) &= \Phi^{-1}(t)\left(-b_x^*(t) + \sum_{j=1}^m \left(\sigma_x^{j,*}(t)\right)^2\right)dt - \sum_{j=1}^m \Phi^{-1}(t)\sigma_x^{j,*}(t)dW_j(t), \\ \Phi^{-1}(0) &= I\,. \end{cases} \tag{3.7}$$

From the Itô's formula we can obtain the solution of the equation (3.5) by using $\Phi(t), \Phi^{-1}(t)$.

$$\begin{aligned} y(t) &= \Phi(t)\int_0^t \Phi^{-1}(s)\left(b_u^*(s) - \sigma_x^*(s)\sigma_u^*(s)\right)v(s)ds \\ &\quad + \sum_{j=1}^m \Phi(t)\int_0^t \Phi^{-1}(s)\sigma_u^{j,*}(s)v(s)dW_j(s)\,, \end{aligned} \tag{3.8}$$

where

$$\sigma_x^*(s)\sigma_u^*(s) = \sum_{j=1}^m \sigma_x^{j,*}(s)\sigma_u^{j,*}(s)\,. \tag{3.9}$$

Since $u^*(\cdot)$ is an optimal control, it is a critical point of the functional $J(u(\cdot))$, namely, we have

$$\frac{d}{d\varepsilon}J(u^*(\cdot) + \varepsilon v(\cdot))\Big|_{\varepsilon=0} = 0\,. \tag{3.10}$$

But

$$\frac{d}{d\varepsilon}J(u^*(\cdot) + \varepsilon v(\cdot))\Big|_{\varepsilon=0} = \mathbb{E}\left(\int_0^T \left(l_x^{*\top}(t)y(t) + l_u^{*\top}(t)v(t)\right)dt\right) + \mathbb{E}\left(h_x^\top(x^*(T))y(T)\right)\,.$$

Using (3.8) for the solution $y(t)$, we have

$$
\begin{aligned}
& \frac{d}{d\varepsilon} J(u^*(\cdot)+\varepsilon v(\cdot))\Big|_{\varepsilon=0} \\
= \;& \mathbb{E}\int_0^T \left(l_x^{*\top}(t)\Phi(t)\left[\int_0^t \Phi^{-1}(s)\left(b_u^*(s)-\sigma_x^*(s)\sigma_u^*(s)\right)v(s)ds\right] + l_u^{*\top}(t)v(t)\right) dt \\
& +\mathbb{E}\int_0^T \left(\sum_{j=1}^m l_x^{*\top}(t)\Phi(t)\left[\int_0^t \Phi^{-1}(s)\sigma_u^{j,*}(s)v(s)dW_j(s)\right]\right) dt \\
& +\mathbb{E}\left\{ h_x^\top(x^*(T))\Phi(T)\left[\int_0^T \Phi^{-1}(s)(b_u^*(s)-\sigma_x^*(s)\sigma_u^*(s))v(s)ds\right]\right\} \\
& +\sum_{j=1}^m \mathbb{E}\left\{ h_x^\top(x^*(T))\Phi(T)\left[\int_0^T \Phi^{-1}(s)\sigma_u^{j,*}(s)v(s)dW_j(s)\right]\right\}.
\end{aligned}
$$

Now we make use of the following identity from Malliavin calculus:

$$
\mathbb{E}\left(F\int_0^T g(t)dW_j(t)\right) = \mathbb{E}\int_0^T \left(D_t^j F\right) g(t)dt\,.
$$

Combining this identity and Fubini type theorem, we have

$$
\begin{aligned}
& \frac{d}{d\varepsilon} J(u^*(\cdot)+\varepsilon v(\cdot))\Big|_{\varepsilon=0} \\
= \;& \mathbb{E}\int_0^T \left(\int_s^T l_x^{*\top}(t)\Phi(t)\Phi^{-1}(s)(b_u^*(s)-\sigma_x^*(s)\sigma_u^*(s))v(s)dt\right) ds \\
& +\int_0^T \mathbb{E}\left(\sum_{j=1}^m \int_0^t D_s^j\left(l_x^{*\top}(t)\Phi(t)\right)\Phi^{-1}(s)\sigma_u^{j,*}(s)v(s)ds\right) dt \\
& +\mathbb{E}\int_0^T l_u^{*\top}(s)v(s)ds + \mathbb{E}\int_0^T h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)(b_u^*(s)-\sigma_x^*(s)\sigma_u^*(s))v(s)ds \\
& +\sum_{j=1}^m \mathbb{E}\int_0^T D_s^j\left(h_x^\top(x^*(T))\Phi(T)\right)\Phi^{-1}(s)\sigma_u^{j,*}(s)v(s)ds\,. \qquad (3.11)
\end{aligned}
$$

Denote

$$
\begin{aligned}
\Psi(T,s) \;:=\; & \left(\int_s^T l_x^{*\top}(t)\Phi(t)dt\right)\Phi^{-1}(s)(b_u^*(s)-\sigma_x^*(s)\sigma_u^*(s)) + l_u^{*\top}(s) \\
& +\sum_{j=1}^m \left(\int_s^T D_s^j\left(l_x^{*\top}(t)\Phi(t)dt\right)\Phi^{-1}(s)\sigma_u^{j,*}(s) + D_s^j\left(h_x^\top(x^*(T))\Phi(T)\right)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right) \\
& +h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)(b_u^*(s)-\sigma_x^*(s)\sigma_u^*(s)). \qquad (3.12)
\end{aligned}
$$

Then from (3.10) and (3.11) we have

$$\begin{aligned} \frac{d}{d\varepsilon}J(u^*(\cdot)+\varepsilon v(\cdot))\Big|_{\varepsilon=0} &= \mathbb{E}\left[\int_0^T \Psi(T,s)v(s)ds\right] \\ &= \mathbb{E}\left[\int_0^T \mathbb{E}\left\{\Psi(T,s)\big|\mathcal{G}_s\right\}v(s)ds\right]=0\,. \end{aligned}$$

Since the above identity holds true for all $(\mathcal{G}_t)_{0\le t\le T}$ adapted process $v\in U[0,T]$, we have

$$\mathbb{E}\left\{\Psi(T,s)\big|\mathcal{G}_s\right\}=0\quad \forall\ 0\le s\le T\,. \tag{3.13}$$

We can also write the above equation as

$$\mathbb{E}\left\{\Psi(T,s)^\top\big|\mathcal{G}_s\right\}=0\quad \forall\ 0\le s\le T\,. \tag{3.14}$$

Denote $\mathbb{E}^{\mathcal{G}_t}(X)=\mathbb{E}\left\{X\big|\mathcal{G}_t\right\}$. Then we can state (3.14) as the following general maximum principle (e.g. the equation (3.15) below).

Theorem 3.2 *Let $(x^*(\cdot),u^*(\cdot))$ be an optimal pair of the control problem (3.1)-(3.3). Define*

$$\begin{cases} P(t) &= \Phi^{\top-1}(t)\int_t^T\Phi^\top(s)l_x^*(s)ds+\Phi^{\top-1}(t)\Phi^\top(T)h_x(x^*(T))\,, \\ Q_j(t) &= -\sigma_x^{j,*}(t)P(t)+\Phi^{\top-1}(t)D_t\left(\Phi^\top(T)h_x(x^*(T))\right)+\int_t^T D_t\left(\Phi^\top(s)l_x^*(s)\right)ds\,. \end{cases}$$

Then

$$\mathbb{E}^{\mathcal{G}_t}\left[b_u^{*\top}(t)P(t)+\sum_{j=1}^m\sigma_u^{j,*\top}(t)Q_j(t)+l_u^*(t)\right]=0\qquad\forall\ \ t\in[0,T] \tag{3.15}$$

almost surely.

3.2 Stochastic control with complete information

If $\mathcal{F}_t=\sigma(W_1(s),\cdots,W_m(s),0\le s\le t)$ is the σ-algebra generated by the Brownian motion $W(s)=(W_1(s),\cdots,W_m(s))$, then the above equation (3.15) for the maximum principle can be simplified.

First note that $b_u^{*\top}(t)$, $\sigma_u^{1,*\top}(t)$, $\cdots$, $\sigma_u^{1,*\top}(t)$, and $l_u^*(t)$ are $\mathcal{F}_t$-adapted, then the equation (3.15) can be written as

$$b_u^{*\top}(t)p(t)+\sum_{j=1}^m\sigma_u^{j,*\top}(t)q_j(t)+l_u^*(t)=0\qquad\forall\ \ t\in[0,T]\,, \tag{3.16}$$

where we denote $p(t)=\mathbb{E}^{\mathcal{F}_t}[P(t)]$ and $q_j(t)=\mathbb{E}^{\mathcal{F}_t}[Q_j(t)]$. From the definition of $P(t)$ and $Q_j(t)$, we have

$$\begin{cases} p(t) &= \Phi^{\top-1}(t)\mathbb{E}^{\mathcal{F}_t}\left[\int_t^T\Phi^\top(s)l_x^*(s)ds+\Phi^\top(T)h_x(x^*(T))\right], \\ q_j(t) &= -\left(\sigma_x^{j,*}(t)\right)p(t)+\Phi^{\top-1}(t)\mathbb{E}^{\mathcal{F}_t}\Big[D_t\left(\Phi^\top(T)h_x(x^*(T))\right) \\ &\qquad +\int_t^T D_t\left(\Phi^\top(s)l_x^*(s)\right)ds\Big]\,,\qquad j=1,\cdots,m. \end{cases} \tag{3.17}$$

Lemma 3.3 *If $p(t)$ and $q_j(t)$ are defined as above, then*

$$q_j(t) = D_t^j p(t)\,. \tag{3.18}$$

Proof From (3.7) we see that

$$D_t^j \Phi^{-1}(t) = -\Phi^{-1}(t)\sigma_x^{j,*}(t)\,.$$

On the other hand, from Proposition 1.2.8 of [27], we see that

$$\mathbb{E}^{\mathcal{F}_t}(D_t^j X) = D_t^j \mathbb{E}^{\mathcal{F}_t}(X)\,.$$

This proves the lemma easily. ■

From the equation (2.11) of [19], we see that $p(t)$ and $(q_1(t)\,, \cdots, q_m(t))$ is the unique solution of the following backward stochastic differential equations.

$$\begin{cases} -dp(t) &= (b_x^{*\top}(t)p(t) + \sum_{j=1}^m \sigma_x^{j,*\top}(t)q_j(t) + l_x^*(t))dt - \sum_{j=1}^m q_j(t)dW_j(t)\,, \\ p(T) &= h_x(x^*(T))\,. \end{cases} \tag{3.19}$$

Theorem 3.4 *Let $(x^*(\cdot), u^*(\cdot))$ be an optimal pair of the control problem (3.1)-(3.3). Let $p(t)$ and $(q_1(t), \cdots, q_m(t))$ be the unique solution pair to (3.19). Then*

$$b_u^{*\top}(t)p(t) + \sum_{j=1}^m \sigma_u^{j,*\top}(t)q_j(t) + l_u^*(t) = 0 \qquad \forall \quad t \in [0,T] \tag{3.20}$$

almost surely.

Remark 3.5 *The equations (3.1), (3.19), and (3.20) is a system of coupled forward-backward stochastic differential equations. Usually they can be used to determine the optimal control u^* and the corresponding optimal state x^*. For the convenience we can write them together as*

$$\begin{cases} dx^*(t) = b(t, x^*(t), u^*(t))dt + \sum_{j=1}^m \sigma^j(t, x^*(t), u^*(t))dW_j(t), \\ x^*(0) = x_0\,, \\ -dp(t) = (b_x^{*\top}(t)p(t) + \sum_{j=1}^m \sigma_x^{j,*\top}(t)q_j(t) + l_x^*(t))dt - \sum_{j=1}^m q_j(t)dW_j(t)\,, \\ p(T) = h_x(x^*(T)) \\ b_u^{*\top}(t)p(t) + \sum_{j=1}^m \sigma_u^{j,*\top}(t)q_j(t) + l_u^*(t) = 0\,, \end{cases} \tag{3.21}$$

where $0 \le t \le T$.

The system of coupled forward-backward stochastic differential equations (3.1), (3.19), and (3.20) can also be written by using the so-called Hamiltonian. Let

$$\begin{aligned} H(t,x,u,p,q) \;&:=\; b(t,x,u)^\top p(t) + \sum_{j=1}^m \sigma^{j,\top}(t,x,u)q_j(t) + l(t,x,u), \\ &\qquad (t,x,u,p,q) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^n \times \mathbb{R}^{n\times m}\,. \end{aligned}$$

Then the maximum principle (3.21) can be restated as

$$\begin{cases} dx^*(t) = \dfrac{\partial}{\partial p} H(t, x^*(t), u^*(t), p(t), q(t))dt + \displaystyle\sum_{j=1}^{m} \dfrac{\partial}{\partial q_j} H(t, x^*(t), u^*(t), p(t), q(t))dW_j(t) \\ x^*(0) = x_0\,, \\ -dp(t) = \dfrac{\partial}{\partial x} H(t, x^*(t), u^*(t), p(t), q(t))dt - \displaystyle\sum_{j=1}^{m} q_j(t)dW_j(t)\,, \\ p(T) = h_x(x^*(T)) \\ \dfrac{\partial}{\partial u} H(t, x^*(t), u^*(t), p(t), q(t)) = 0\,, \end{cases} \tag{3.22}$$

where $0 \le t \le T$.

4 Controlled stochastic differential systems driven by fractional Brownian motions

To obtain our main results of maximum principle for the system driven by fractional Brownian motions, we need to develop some new results for controlled stochastic differential equations driven by fractional Brownian motions. Let us recall that $B^H(t) = (B_1^H(t), \cdots, B_m^H(t)), 0 \le t \le T$, is an m-dimensional Brownian motion. Let $(\mathcal{G}_t\,, 0 \le t \le T)$ be a right continuous filtration contained in the filtration $(\mathcal{F}_t\,, 0 \le t \le T)$ generated by fractional Brownian motions $B^H(t)$.

First, let us define our space of admissible controls:

$$\begin{aligned} U[0,T] \quad := \quad \Big\{ & u | u : [0,T] \times \Omega \to \mathbb{R}^d, \mathcal{G}\text{-adapted}, u \in C^\mu[0,T] \text{ for some } \mu > 1 - H\,, \\ & \text{there exist constant } C > 0,\ c > 0, \text{ and } \beta < H \\ & \quad \text{such that } \ \|u\|_{\mu,0,T} \le Ce^{c\sum_{j=1}^m \|B_j^H\|_{\beta,0,T}} \\ & \quad \text{and } \ \int_0^T \int_0^T \mathbb{E}\left(|D_s^H u(t)|^2\right) dsdt < \infty \Big\}. \end{aligned} \tag{4.1}$$

Consider the following controlled stochastic differential equation driven by fractional Brownian motion:

$$\begin{cases} dx(t) &= b(t, x(t), u(t))dt + \sum_{j=1}^{m} \sigma^j(t, x(t), u(t))d^\circ B_j^H(t), \\ x(0) &= x_0\,. \end{cases} \tag{4.2}$$

Here the integral with respect to fractional Brownian motion is in the pathwise sense (or the Stratonovich type integral).

We assume that $b : [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$, $\sigma : [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^{n\times m}$ are some given continuous functions satisfying the following conditions.

(H4) $b(t,x,u)$ is continuously differentiable with respect to x and u. Moreover, there exists a constant L such that the following properties hold:

$$
\begin{aligned}
& |b_x(t,x,u)| + |b_u(t,x,u)| \le L\,, \\
& |b_x(t,x,u) - b_x(t,y,v)| + |b_u(t,x,u) - b_u(t,y,v)| \le L(|x-y| + |u-v|)\,, \\
& \qquad \forall x \ \in \mathbb{R}^n, \ \ \forall \ \ u \in \mathbb{R}^d, \ \ \forall t \in [0,T]\,.
\end{aligned} \tag{4.3}
$$

(H5) $\sigma(t,x,u)$ is twice continuously differentiable in x and u and there exist some constants $1-H < \gamma < 1$ and $0 < \delta \le 1$, and L such that for each $i = 1, \cdots, n$:

(i) The partial derivatives of σ with respect to x and u are bounded:

$$
\begin{aligned}
|\sigma_x(t,x,u)| + |\sigma_u(t,x,u)| \quad &\le \quad L\,, \\
|\sigma_{xx}(t,x,u)| + |\sigma_{uu}(t,x,u)| + \sigma_{xu}(t,x,u)| \quad &\le \quad L\,.
\end{aligned}
$$

(ii) Hölder continuity in time: $\forall x \in \mathbb{R}^n, u \in \mathbb{R}, \forall t,s \in [0,T]$,

$$
\begin{aligned}
& |\sigma(t,x,u) - \sigma(s,x,u)| + |\partial_x \sigma(t,x,u) - \partial_x \sigma(s,x,u)| \\
& + |\partial_u \sigma(t,x,u) - \partial_u \sigma(s,x,u)| + |\sigma_{xx}(t,x,u) - \sigma_{xx}(s,x,u)| \\
& + |\sigma_{xu}(t,x,u) - \sigma_{xu}(s,x,u)| + |\sigma_{uu}(t,x,u) - \sigma_{uu}(s,x,u)| \le \quad L|t-s|^{\gamma}.
\end{aligned}
$$

(iii) Lipschitz continuity of second derivatives with respect to state and control variables.

$$
\begin{aligned}
& |\sigma_{uu}(t,x,u) - \sigma_{uu}(t,y,v)| + |\sigma_{xu}(t,x,u) - \sigma_{xu}(t,y,v)| \\
\le \quad & L\left(|x-y| + |u-v|\right)\,.
\end{aligned}
$$

If b and σ satisfy the above assumptions (H4) and (H5) and if u is an admissible control, then the coefficients $b(t,x) = b(t,u(t),x)$ and $\sigma(t,x) = \sigma(t,u(t),x)$ satisfy the conditions (H_1), (H_2), and (H_3) of [28]. Thus it follows that for any admissible control u, the controlled stochastic differential equation (4.2) has a unique solution (see also [17]), denoted by x_t^u. Moreover, for any $1-H < \alpha < 1/2$, the solution is $1-\alpha$ Hölder continuous almost surely, namely, $|x(r) - x(\tau)| \le c_0 |r-\tau|^{1-\alpha}$ almost surely (where c_0 may depends on $B_{\cdot}^H$). The solution x_t^u of above equation (4.2) depends on u. But to simplify the notation we often omit its explicit dependence on u and write $x_t = x_t^u$.

Let u^* and v be two admissible controls. Denote $\bar{u} = v - u^*$. For any $\varepsilon \in \mathbb{R}$, we denote $u^\varepsilon = u^* + \varepsilon \bar{u}$. Then $\bar{u}$ and u^ε are also admissible controls. Corresponding to u^* and u^ε there are solutions $x^\varepsilon(\cdot) = x(\cdot; u^\varepsilon(\cdot))$ and $x^*(\cdot) =$

$x(\cdot; u^*(\cdot))$ to the equations (4.2) . That is

$$x^*(t) = x_0 + \int_0^t b(s, x^*(s), u^*(s))ds + \sum_{j=1}^m \int_0^t \sigma_j(s, x^*(s), u^*(s))dB_j^H(s), \tag{4.4}$$

$$x^\varepsilon(t) = x_0 + \int_0^t b(s, x^\varepsilon(s), u^\varepsilon(s))ds + \sum_{j=1}^m \int_0^t \sigma_j(s, x^\varepsilon(s), u^\varepsilon(s))dB_j^H(s) . \tag{4.5}$$

To obtain our results of maximum principle, we need the following.

Proposition 4.1 *Assume (H4) and (H5). Let $x^\varepsilon(\cdot)$ and $x^*(\cdot)$ be the solutions of equation (4.2) corresponding to $u^\varepsilon(\cdot)$ and $u^*(\cdot)$ respectively. Then*

$$\begin{aligned} \lim_{\varepsilon\to 0} \sup_{0\le t\le T} |x^\varepsilon(t) - x^*(t)| &= 0 , \\ \lim_{\varepsilon\to 0} \sup_{0\le t\le T} \left| \frac{x^\varepsilon(t) - x^*(t)}{\varepsilon} - y(t) \right| &= 0 \end{aligned}$$

almost surely as $\varepsilon \to 0$, where $y(t)$ is the solution of the following equation.

$$\begin{aligned} y(t) &= \int_0^t \left[b_x^*(s)y(s) + b_u^*(s)\bar{u}(s) \right] ds \\ &\quad + \int_0^t \sum_{j=1}^m \left[\sigma_x^{j*}(s)y(s) + \sigma_u^{j*}(s)\bar{u}(s) \right] d^\circ B_j^H(s) . \end{aligned} \tag{4.6}$$

Proof of Proposition 4.1 We shall follow the idea of [17] to prove the above lemma. Fix $\frac{1}{2} > \alpha > 0$ and $\beta > 0$ such that $0 < 1 - \alpha < \beta < H$. Let $g_1(t), \cdots, g_m(t)$ be any given functions of β- Hölder continuous. Consider the following deterministic differential equation

$$x(t) = x_0 + \int_0^t b(s, x(s), u(s))ds + \sum_{j=1}^m \int_0^t \sigma_j(s, x(s), u(s))dg_j(s) . \tag{4.7}$$

Corresponding to the admissible controls u^* and u^ε, the above equation (4.7) has also two solutions, still denoted by x^* and x^ε. Corresponding to u^ε, we shall obtain an equations depending on a parameter ε. We can consider more general one:

$$x^\varepsilon(t) = x_0 + \int_0^t b(\varepsilon, s, x^\varepsilon(s))ds + \sum_{j=1}^m \int_0^t \sigma_j(\varepsilon, s, x^\varepsilon(s))dg_j(s) . \tag{4.8}$$

Lemma 4.2 *Let $b(\varepsilon, t, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma_j(\varepsilon, t, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$, $j = 1, \cdots, m$, be a continuously differentiable functions with uniformly bounded derivatives.*

Let $\sigma_j(\varepsilon,\cdot,x):[0,T]\to\mathbb{R}^n$, $j=1,\cdots,m$, be uniformly Hölder continuous with exponent $\gamma>\alpha$. That means that there is a constant $M\in(0,\infty)$, independent of ε, t, and x, such that

$$\left|\frac{\partial}{\partial x_i}b_j(\varepsilon,t,x)\right|\le M\,,\quad \left|\frac{\partial}{\partial x_i}\sigma_j(\varepsilon,t,x)\right|\le M$$

and

$$|\sigma_j(\varepsilon,t,x)-\sigma_j(\varepsilon,s,x)|\le M|t-s|^{\gamma}\,.$$

Assume also that b and σ_j satisfy the following uniform linear growth condition:

$$|b_j(\varepsilon,t,x)|\le M(1+|x|)\,,\quad |\,\sigma_j(\varepsilon,t,x)|\le M(1+|x|)\,.$$

Then there are constants C and c independent of ε, M, and g, such that for all T,

$$\sup_{0\le t\le T}|x^\varepsilon(t)|\ \le Ce^{c\|g\|_{0,T,\beta}^{\frac{1}{\beta}}}(|x_0|+1) \tag{4.9}$$

and

$$\|x^\varepsilon\|_{0,T,\beta}\le Ce^{c\|g\|_{0,T,\beta}^{\frac{1}{\beta}}}(|x_0|+1)\,. \tag{4.10}$$

Remark 4.3 *We also have that $\mathbb{E}\sup_{0\le t\le T}|x^\varepsilon(t)|^p$ and $\mathbb{E}\|x^\varepsilon\|_{0,T,\beta}^p$ for any $p>0$ are uniformly bounded (independent of ε), when $g=B^H$ and $\frac{1}{2}<\beta<H$, by Fernique theorem. Actually, since $\frac{1}{\beta}<2$, then by Fernique theorem, we have $\mathbb{E}e^{p\|B^H\|_{0,T,\beta}^{\frac{1}{\beta}}}<\infty$ for all $p>0$.*

Proof The existence of the solution x^ε for every ε is known. See example in [17] and [28]. We shall sketch the proof of the bounds (4.9) following idea in the proof of Theorem 3.1 of [17]. Without loss of generality we assume that $n=m=1$. Set $\|g\|_\beta=\|g\|_{0,T,\beta}$. We can assume that $\|g\|_\beta>0$, otherwise the inequalities are obvious.

Step 1. From (2.5), we have for any $0<s<t<T$

$$\begin{aligned}\left|\int_s^t\sigma(\varepsilon,r,x^\varepsilon(r))dg_r\right| &\le C\|g\|_\beta\left[\int_s^t\frac{|\sigma(\varepsilon,r,x^\varepsilon(r))|}{(r-s)^\alpha}(t-r)^{\alpha+\beta-1}dr\right.\\ &\quad\left.+\int_s^t\int_s^r\frac{|\sigma(\varepsilon,r,x^\varepsilon(r))-\sigma(\varepsilon,\tau,x^\varepsilon(\tau))|}{|r-\tau|^{\alpha+1}}(t-r)^{\alpha+\beta-1}d\tau dr\right]\\ =:\quad I_1+I_2\,,\end{aligned} \tag{4.11}$$

where and in what follows, C is a universal constant (independent g and M). It is easy to see from the assumption of the lemma that

$$I_1\le CM\|g\|_\beta\left[1+\|x^\varepsilon\|_{s,t,\infty}\right](t-s)^\beta\,. \tag{4.12}$$

I_2 can be estimated as

$$\begin{aligned} I_2 &\leq C\|g\|_\beta \left[\int_s^t \int_s^r \frac{|\sigma(\varepsilon,r,x^\varepsilon(r))-\sigma(\varepsilon,\tau,x^\varepsilon(r))|}{|r-\tau|^{\alpha+1}}(t-r)^{\alpha+\beta-1}d\tau dr \right.\\ &\qquad \left. \int_s^t \int_s^r \frac{|\sigma(\varepsilon,\tau,x^\varepsilon(r))-\sigma(\varepsilon,\tau,x^\varepsilon(\tau))|}{|r-\tau|^{\alpha+1}}(t-r)^{\alpha+\beta-1}d\tau dr \right]\\ &\leq CM\|g\|_\beta \left[(t-s)^\gamma + \|x^\varepsilon\|_{s,t,\beta}(t-s)^\beta\right](t-s)^\beta . \end{aligned} \tag{4.13}$$

On the other hand

$$\left|\int_s^t b(\varepsilon,r,x^\varepsilon(r))dr\right| \leq CM\left(1+\|x^\varepsilon\|_{s,t,\infty}\right)(t-s). \tag{4.14}$$

Therefore from (4.11)-(4.14) we see that the solution x^ε to (4.8) satisfies

$$\|x^\varepsilon\|_{s,t,\beta} \leq CM\left[1+\|x^\varepsilon\|_{s,t,\infty}\right] + CM\|g\|_\beta\left[1+\|x^\varepsilon\|_{s,t,\infty}+\|x^\varepsilon\|_{s,t,\beta}(t-s)^\beta\right], \text{ for } s,t\in[0,T]. \tag{4.15}$$

Step 2. Choose Δ such that

$$\Delta = \left(\frac{1}{3CM[1+\|g\|_\beta]}\right)^{\frac{1}{\beta}}. \tag{4.16}$$

Then, for all s and t such that $0\leq t-s\leq \Delta$ we have

$$\|x^\varepsilon\|_{s,t,\beta} \leq \frac{3}{2}CM[1+\|g\|_\beta]\left(1+\|x^\varepsilon\|_{s,t,\infty}\right). \tag{4.17}$$

Therefore, when $0\leq t-s\leq\Delta$

$$|x^\varepsilon(t)| \leq |x^\varepsilon(s)| + \frac{3}{2}CM[1+\|g\|_\beta]\left(1+\|x^\varepsilon\|_{s,t,\infty}\right)\Delta^\beta, \tag{4.18}$$

or

$$\|x^\varepsilon\|_{s,t,\infty} \leq |x^\varepsilon(s)| + \frac{3}{2}CM[1+\|g\|_\beta]\left(1+\|x^\varepsilon\|_{s,t,\infty}\right)\Delta^\beta$$

for $0\leq t-s\leq\Delta$. Using again (4.16) we get

$$\|x^\varepsilon\|_{s,t,\infty} \leq 2|x^\varepsilon(s)| + 3CM[1+\|g\|_\beta]\Delta^\beta.$$

Since (4.16) implies

$$\Delta \leq \left(\frac{2}{3CM[1+\|g\|_\beta]}\right)^{\frac{1}{\beta}}.$$

Then

$$\|x^\varepsilon\|_{s,t,\infty} \leq 2\left(|x^\varepsilon(s)|+1\right).$$

Hence,

$$\sup_{0\leq r\leq t}|x^\varepsilon(r)| \leq 2\left(\sup_{0\leq r\leq s}|x^\varepsilon(r)|+1\right) \quad \forall\, t-s\leq\Delta\,, 0\leq s\leq t\leq T. \tag{4.19}$$

Now we divide the interval $[0, T]$ into $n = [T/\Delta] + 1$ subintervals, and use the estimate (4.19) in every interval to obtain

$$\sup_{0 \le t \le T} |x^\varepsilon(t)| \le 2^n \left(|x^\varepsilon(0)| + 1\right) \le 2^{\frac{T}{\Delta}+1} \left(|x^\varepsilon(0)| + 1\right) \ .$$

Finally, we have from (4.16)

$$\sup_{0 \le t \le T} |x^\varepsilon(t)| \ \le Ce^{c\|g\|_{0,T,\beta}^{\frac{1}{\beta}}} (|x^\varepsilon(0)| + 1) \,.$$

Step 3. From Equation (4.17), we see also when $t - s \le \Delta$ and $0 \le s < t \le T$

$$\|x^\varepsilon\|_{s,t,\beta} \ \le Ce^{cTM^{\frac{1}{\beta}} \|g\|_{0,T,\beta}^{\frac{1}{\beta}}} (|x^\varepsilon(0)| + 1) \tag{4.20}$$

since $x \le Ce^{cx^{1/\beta}}, \forall \ x \ge 0$ with some constants c and C. For general $0 \le s < t \le T$, we denote $s = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$ so that $t_k - t_{k-1} \le \Delta$. Then

$$\begin{aligned}
\left|\frac{x^\varepsilon(t) - x^\varepsilon(s)}{(t-s)^\beta}\right| &\le \sum_{k=0}^{n-1} \left|\frac{x^\varepsilon(t_{k+1}) - x^\varepsilon(t_k)}{(t_{k+1} - t_k)^\beta}\right| \\
&\le \sum_{k=0}^{n-1} Ce^{c\|g\|_{0,T,\beta}^{\frac{1}{\beta}}} (|x^\varepsilon(0)| + 1) \\
&\le nCe^{c\|g\|_{0,T,\beta}^{\frac{1}{\beta}}} (|x^\varepsilon(0)| + 1) \,.
\end{aligned}$$

With the same argument as (4.20), we have (4.10). The proof of the theorem is now complete. ■

It is clear that the Proposition 4.1 is a consequence of the following lemma.

Lemma 4.4 *Assume (H4) and (H5) and assume* $0 < 1 - \mu < 1 - \alpha < \beta < H$. *Let* $g_1(t), \cdots, g_m(t)$ *be* β*-Hölder continuous functions of* $t \in [0, T]$. *Let* $x^\varepsilon(\cdot)$ *and* $x^*(\cdot)$ *be the solutions of equations (4.7) corresponding to* $u^\varepsilon(\cdot)$ *and* $u^*(\cdot)$, *respectively. Then*

$$\begin{gathered}
\lim_{\varepsilon \to 0} \sup_{0 \le t \le T} |x^\varepsilon(t) - x^*(t)| = 0 \,, \\
\lim_{\varepsilon \to 0} \|x^\varepsilon - x^*\|_{1-\alpha} = 0 \,, \\
\lim_{\varepsilon \to 0} \sup_{0 \le t \le T} \left|\frac{x^\varepsilon(t) - x^*(t)}{\varepsilon} - y(t)\right| = 0 \,,
\end{gathered}$$

where $y(t)$ *is the solution of the following linear equation.*

$$y(t) = \int_0^t [b_x^*(s)y(s) + b_u^*(s)\bar{u}(s)]ds + \int_0^t \sum_{j=1}^m [\sigma_x^{j,*}(s)y(s) + \sigma_u^{j,*}(s)\bar{u}(s)]dg_j(s). \tag{4.21}$$

Moreover, the above limits hold in L^p *sense as well for all* $p > 0$.

Proof We follow the idea of [17]. We divide the proof into several steps.

Step 1. To simplify the notation, we assume $n = d = m = 1$. The general case only increases the notational complexity. Throughout this paper we shall use C to denote a generic constant, independent of ε, whose values may be different in different occurrences. Denote $\sigma(\cdot, x^\varepsilon(\cdot), u^\varepsilon(\cdot))$ and $\sigma(\cdot, x^*(\cdot), u^*(\cdot))$ by $\sigma^\varepsilon(\cdot)$ and $\sigma^*(\cdot)$ respectively. Set $y^\varepsilon(\cdot) = x^\varepsilon(\cdot) - x^*(\cdot)$. We have that

$$\begin{aligned} |y^\varepsilon(t) - y^\varepsilon(s)| &\leq \left|\int_s^t (b^\varepsilon(r) - b^*(r))\, dr\right| + \left|\int_s^t [\sigma^\varepsilon(r) - \sigma^*(r)] dg(r)\right| \\ &\leq L\int_s^t |y^\varepsilon(r)| dr + \varepsilon L \int_s^t |\bar{u}(r)| dr \\ &\qquad + k\|g\|_\beta \int_s^t (t-r)^{\alpha+\beta-1} \left[D_{s+}^\alpha (\sigma^\varepsilon(r) - \sigma^*(r)\right] dr \\ &\leq L\int_s^t |y^\varepsilon(r)| dr + \varepsilon L \int_s^t |\bar{u}(r)| dr + I_1 + I_2\,, \end{aligned} \tag{4.22}$$

where

$$\begin{aligned} I_1 &:= k\|g\|_\beta \int_s^t (t-r)^{\alpha+\beta-1} \frac{|\sigma^\varepsilon(r) - \sigma^*(r)|}{(r-s)^\alpha} dr \\ I_2 &:= k\|g\|_\beta \int_s^t \int_s^r (t-r)^{\alpha+\beta-1} \frac{|\sigma^\varepsilon(r) - \sigma^*(r) - [\sigma^\varepsilon(\tau) - \sigma^*(\tau)]|}{(r-\tau)^{\alpha+1}} d\tau dr\,. \end{aligned}$$

I_1 is handled easily.

$$\begin{aligned} I_1 &\leq C\|g\|_\beta \int_s^t \frac{|y^\varepsilon(r)|(t-r)^{\alpha+\beta-1}}{(r-s)^\alpha} dr + C\varepsilon \|g\|_\beta \int_s^t \frac{|\bar{u}(r)|(t-r)^{\alpha+\beta-1}}{(r-s)^\alpha} dr\,. \\ &\leq C\|g\|_\beta (t-s)^\beta \sup_{s\leq r\leq t} |y^\varepsilon(r)| + C\varepsilon\|g\|_\beta (t-s)^\beta\,, \end{aligned} \tag{4.23}$$

Step 2. To estimate I_2, let us consider the integral in I_2, denoted by $\tilde{I}_2$. .

$$\begin{aligned}
\tilde{I}_2 &:= \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} \left|\sigma^\varepsilon(r) - \sigma^*(r) - [\sigma^\varepsilon(\tau) - \sigma^*(\tau)]\right| d\tau dr \\
&= \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} \Big| \int_0^1 \sigma_x(r, x^*(r) + \lambda y^\varepsilon(r), u^\varepsilon(r)) d\lambda \cdot y^\varepsilon(r) \\
&\quad + \int_0^1 \sigma_u(r, x^*(r), u^*(r) + \lambda(u^\varepsilon(r) - u^*(r))) d\lambda \cdot (u^\varepsilon(r) - u^*(r)) \\
&\quad - \int_0^1 \sigma_x(\tau, x^*(\tau) + \lambda y^\varepsilon(\tau), u^\varepsilon(\tau)) d\lambda \cdot y^\varepsilon(\tau) \\
&\quad - \int_0^1 \sigma_u(\tau, x^*(\tau), u^*(\tau) + \lambda(u^\varepsilon(\tau) - u^*(\tau))) d\lambda \cdot (u^\varepsilon(\tau) - u^*(\tau)) \Big| d\tau dr \\
&= \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} \Big| \int_0^1 \sigma_x(r, x^*(r) + \lambda y^\varepsilon(r), u^\varepsilon(r)) d\lambda \cdot [y^\varepsilon(r) - y^\varepsilon(\tau)] \\
&\quad + \int_0^1 [\sigma_x(r, x^*(r) + \lambda y^\varepsilon(r), u^\varepsilon(r)) - \sigma_x(\tau, x^*(\tau) + \lambda y^\varepsilon(\tau), u^\varepsilon(\tau))] d\lambda \cdot y^\varepsilon(\tau) \\
&\quad + \int_0^1 \sigma_u(r, x^*(r), u^*(r) + \lambda\varepsilon\bar{u}(r)) d\lambda \cdot \varepsilon(\bar{u}(r) - \bar{u}(\tau)) \\
&\quad + \int_0^1 [\sigma_u(r, x^*(r), u^*(r) + \lambda\varepsilon\bar{u}(r)) - \sigma_u(\tau, x^*(\tau), u^*(\tau) + \lambda\varepsilon\bar{u}(\tau))] d\lambda \cdot \varepsilon\bar{u}(\tau) \Big| d\tau dr\,.
\end{aligned}$$

Since σ_x and σ_u are bounded, we have

$$\begin{aligned}
\tilde{I}_2 &\leq C \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} |y^\varepsilon(r) - y^\varepsilon(\tau)| d\tau dr \\
&\quad + C\varepsilon \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} |\bar{u}(r) - \bar{u}(\tau)| d\tau dr \\
&\quad + \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} \int_0^1 \Big[|\sigma_x(r, x^*(r) + \lambda y^\varepsilon(r), u^\varepsilon(r)) - \sigma_x(\tau, x^*(r) + \lambda y^\varepsilon(r), u^\varepsilon(r))| \\
&\quad + |\sigma_x(\tau, x^*(r) + \lambda y^\varepsilon(r), u^\varepsilon(r)) - \sigma_x(\tau, x^*(\tau) + \lambda y^\varepsilon(\tau), u^\varepsilon(\tau))| \Big] d\lambda \cdot |y^\varepsilon(\tau)| d\tau dr \\
&\quad + \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} \int_0^1 \Big[|\sigma_u(r, x^*(r), u^*(r) + \lambda\varepsilon\bar{u}(r)) - \sigma_u(\tau, x^*(r), u^*(r) + \lambda\varepsilon\bar{u}(r))| \\
&\quad + |\sigma_u(\tau, x^*(r), u^*(r) + \lambda\varepsilon\bar{u}(r)) - \sigma_u(\tau, x^*(\tau), u^*(\tau) + \lambda\varepsilon\bar{u}(\tau))| \Big] d\lambda \cdot \varepsilon|\bar{u}(\tau)| d\tau dr\,.
\end{aligned}$$

Again since the second derivatives σ_{xx}, σ_{uu}, and σ_{xu} are bounded and the first

derivatives σ_x and σ_u are γ-Hölder continuous in t, we have

$$
\begin{aligned}
\tilde{I}_2 \quad \leq \quad & C\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}|y^\varepsilon(r)-y^\varepsilon(\tau)|d\tau dr \\
& \quad +C\varepsilon\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}|\bar{u}(r)-\bar{u}(\tau)|d\tau dr \\
& +C\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\cdot[(r-\tau)^\gamma+|x^*(r)-x^*(\tau)| \\
& +|x^\varepsilon(r)-x^\varepsilon(\tau)|+|u^\varepsilon(r)-u^\varepsilon(\tau)|]\cdot|y^\varepsilon(\tau)|d\tau dr \\
& +C\varepsilon\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\cdot[(r-\tau)^\gamma+|x^*(r)-x^*(\tau)| \\
& +|u^*(r)-u^*(\tau)|+|u^\varepsilon(r)-u^\varepsilon(\tau)|]\cdot|\bar{u}(\tau)|d\tau dr\,.
\end{aligned}
$$

Since both u and v are admissible controls, they are Hölder continuous of order (μ. Thus u^ε is uniformly Hölder continuous of order μ. From Lemma 4.2, we know that x^ε is also uniformly Hölder continuous of order β and hence x^ε is also uniformly Hölder continuous of order $1-\alpha$

We also use the boundedness of u. Hence, we have

$$
\begin{aligned}
\tilde{I}_2 \quad \leq \quad & C\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}|y^\varepsilon(r)-y^\varepsilon(\tau)|d\tau dr \\
& +C\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}[(r-\tau)^\gamma+(r-\tau)^{(1-\alpha)}+(r-\tau)^\mu]|y^\varepsilon(\tau)|d\tau dr \\
& +C\varepsilon\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}[(r-\tau)^\gamma+(r-\tau)^{(1-\alpha)}+(r-\tau)^\mu]d\tau dr \\
\leq \quad & C\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{2\alpha}}\|y^\varepsilon\|_{1-\alpha,\tau,r}d\tau dr \\
& +C\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}[(r-\tau)^\gamma+(r-\tau)^{(1-\alpha)}+(r-\tau)^\mu]|y^\varepsilon(\tau)|d\tau dr \\
& +C\varepsilon\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}[(r-\tau)^\gamma+(r-\tau)^{(1-\alpha)}+(r-\tau)^\mu]d\tau dr\,.
\end{aligned}
$$

Denote $\nu=(1-\alpha)\wedge\gamma\wedge\mu$. The above inequality gives

$$
\tilde{I}_2 \quad \leq \quad C(t-s)^{1+\beta-\alpha}\|y^\varepsilon\|_{1-\alpha,s,t}+C(t-s)^{\beta+\nu}\|y^\varepsilon\|_{s,t,\infty}+C\varepsilon(t-s)^{\beta+\nu}
$$

Therefore, we have

$$
I_2\leq C\|g\|_\beta\left[(t-s)^{1+\beta-\alpha}\|y^\varepsilon\|_{1-\alpha,s,t}+(t-s)^{\beta+\nu}\|y^\varepsilon\|_{s,t,\infty}+\varepsilon(t-s)^{\beta+\nu}\right]. \tag{4.24}
$$

Combining (4.23) and (4.24), we have

$$
\begin{aligned}
|y^\varepsilon(t)-y^\varepsilon(s)| \quad \leq \quad & C\|g\|_\beta\bigg[(t-s)^{1+\beta-\alpha}\|y^\varepsilon\|_{1-\alpha,s,t}+(t-s)^\beta\|y^\varepsilon\|_{s,t,\infty} \\
& \qquad +\varepsilon(t-s)^\beta\bigg]\,.
\end{aligned} \tag{4.25}
$$

Thus we have

$$\begin{aligned}\|y^\varepsilon\|_{1-\alpha,s,t} \quad &\le \quad C\|g\|_\beta\Big[\|y^\varepsilon\|_{1-\alpha,s,t}(t-s)^\beta + (t-s)^{\beta+\alpha-1}\|y^\varepsilon\|_{s,t,\infty} \\ &\qquad +\varepsilon(t-s)^{\beta+\alpha-1}\Big]. \end{aligned} \tag{4.26}$$

Step 3. We choose Δ_1 such that

$$C\|g\|_\beta \Delta_1^\beta = \frac{1}{2}.$$

From the above equation (4.26) it follows that if $0 < t - s \le \Delta_1$, then

$$\|y^\varepsilon\|_{1-\alpha,s,t} \le C\|g\|_\beta (t-s)^{\alpha+\beta-1}\|y^\varepsilon\|_{s,t,\infty} + C\varepsilon\|g\|_\beta (t-s)^{\alpha+\beta-1}.$$

Since $|y(t)| \le |y(s)| + |t-s|^{1-\alpha}\|y\|_{1-\alpha,s,t}$, the above inequality yields

$$|y^\varepsilon(t)| \le |y^\varepsilon(s)| + C\|g\|_\beta (t-s)^\beta \|y^\varepsilon\|_{s,t,\infty} + C\varepsilon\|g\|_\beta (t-s)^{\ \beta}, \quad \forall\, 0 < t-s \le \Delta_1.$$

which implies easily

$$\|y^\varepsilon\|_{s,t,\infty} \le |y^\varepsilon(s)| + C\|g\|_\beta (t-s)^\beta \|y^\varepsilon\|_{s,t,\infty} + C\varepsilon\|g\|_\beta (t-s)^{\ \beta}, \quad \forall\, 0 < t-s \le \Delta_1. \tag{4.27}$$

Now we choose Δ_2 such that

$$C\|g\|_\beta \Delta_2^\beta = 1/2.$$

[Notice that the constant C may be different than that in the definition for Δ_1.] Then for all $0 \le s < t \le T, t-s \le \Delta_0 := \Delta_1 \wedge \Delta_2$, we have from equation (4.27),

$$|y^\varepsilon|_{\infty,s,t} \le 2|y^\varepsilon(s)| + C\varepsilon\|g\|_\beta (t-s)^\beta. \tag{4.28}$$

We apply the above inequality to $s = 0$ and $t-s \le \Delta_0$ and notice that $y^\varepsilon(0) = 0$. We have that

$$|y^\varepsilon|_{0,\Delta_0,\infty} \le C\varepsilon\|g\|_\beta \Delta_0^\beta.$$

In general, for any integer positive k such that $k\Delta_0 < T$, if we let $s = \Delta_0$ and $t \in [k\Delta_0, (k+1)\Delta_0]$, then we have

$$|y^\varepsilon|_{0,(k+1)\Delta_0,\infty} \le 2|y^\varepsilon_{k\Delta_0}| + C\varepsilon\|g\|_\beta \Delta_0^\beta.$$

This implies

$$|y^\varepsilon|_{0,k\Delta_0,\infty} \le C(2^k - 1)\varepsilon\|g\|_\beta \Delta_0^\beta. \tag{4.29}$$

In the equation (4.29), if we let

$$k = [T/\Delta_0] + 1 \le 2T/\Delta_0 \le 2T/\Delta_1 + 2T/\Delta_2 = CT\|g\|_\beta^{1/\beta}.$$

Then (4.29) yields

$$|y^\varepsilon|_{0,T,\infty} \le C2^{CT\|g\|_\beta^{1/\beta}}\varepsilon. \tag{4.30}$$

Therefore we obtain

$$\lim_{\varepsilon\to 0}\sup_{0\le t\le T}|x^\varepsilon(t)-x^*(t)|=0\,. \tag{4.31}$$

In the same way as in Step 3 in Lemma 4.2, we can also prove

$$\lim_{\varepsilon\to 0}\|x^\varepsilon-x^*\|_{1-\alpha}=0.$$

Step 4. Denote $\eta^\varepsilon(t)=\frac{1}{\varepsilon}y^\varepsilon(t)-y(t)$. Then we can write for all $0\le t\le T$,

$$\begin{aligned}\eta^\varepsilon(t)=&\frac{1}{\varepsilon}\int_0^t[b^\varepsilon(r)-b^*(r)-\varepsilon(b_x^*(r)y(r)+b_u^*(r)\bar u(r))]dr\\ &+\frac{1}{\varepsilon}\int_0^t[\sigma^\varepsilon(r)-\sigma^*(r)-\varepsilon(\sigma_x^*(r)y(r)+\sigma_u^*(r)\bar u(r))]dg(r)\,.\end{aligned}$$

Hence we have

$$|\eta^\varepsilon(t)-\eta^\varepsilon(s)|\quad\le\quad I_3+I_4\,, \tag{4.32}$$

where

$$\begin{aligned}I_3&=\left|\frac{1}{\varepsilon}\int_s^t[b^\varepsilon(r)-b^*(r)-\varepsilon(b_x^*(r)y(r)+b_u^*(r)\bar u(r))]dr\right|\,,\\ I_4&=\left|\frac{1}{\varepsilon}\int_s^t[\sigma^\varepsilon(r)-\sigma^*(r)-\varepsilon(\sigma_x^*(r)y(r)+\sigma_u^*(r)\bar u(r))]dg(r)\right|\,.\end{aligned}$$

Using the argument as for I_1 and from the boundedness and the Lipschitz continuity of the derivative b_x and b_u, and from the inequality (4.30), we have for $0\le s<t\le T$,

$$\begin{aligned}I_3 &= \left|\frac{1}{\varepsilon}\int_s^t\int_0^1\left\{[b_x(r,x^*(r)+\lambda y^\varepsilon(r)\,,u^\varepsilon(r))y^\varepsilon(r)-\varepsilon b_x^*(r)y(r)]\right.\right.\\ &\qquad\left.\left.+\varepsilon\,[b_u(r,x^*(r)\,,u^*(r)+\varepsilon\lambda\bar u(r))\bar u(r)-b_u^*(r)\bar u(r)]\right\}d\lambda dr\right|\\ &\le \int_s^t\int_0^1\left\{|b_x(r,x^*(r)+\lambda y^\varepsilon(r)\,,u^\varepsilon(r))|\,|\eta^\varepsilon(r)|\right.\\ &\qquad+|b_x(r,x^*(r)+\lambda y^\varepsilon(r)\,,u^\varepsilon(r))-b_x^*(r)|\,|y(r)|\\ &\qquad\left.+|b_u(r,x^*(r)\,,u^*(r)+\varepsilon\lambda\bar u(r))-b_u^*(r)|\,|\bar u(r)|\right\}d\lambda dr\Big|\\ &\le C\int_s^t|\eta^\varepsilon(r)|dr+L\int_s^t[(|y^\varepsilon(r)|+\varepsilon|\bar u(r)|)\cdot|y(r)|]\,dr+\varepsilon L\int_s^t|\bar u(r)|^2dr\end{aligned}$$

Since $\sup_{0\le r\le T}|y^\varepsilon|\le C\varepsilon$ we obtain

$$I_3\quad\le\quad C(t-s)\,[\|\eta^\varepsilon\|_{s,t,\infty}+\varepsilon]\,. \tag{4.33}$$

Step 5. To estimate I_4 we shall use the consequence (2.5) of the fractional integration by parts formula (2.4). Denote

$$\bar{\sigma}_\varepsilon(\cdot) = \sigma^\varepsilon(\cdot) - \sigma^*(\cdot) - \varepsilon(\sigma^*_x(\cdot)y(\cdot) + \sigma^*_u(\cdot)\bar{u}(\cdot)).$$

From (2.5) it follows

$$\begin{aligned} I_4 &\leq C\|g\|_\beta \frac{1}{\varepsilon}\int_s^t \frac{|\bar{\sigma}_\varepsilon(r)|}{(t-r)^{1-\alpha-\beta}(r-s)^\alpha}dr \\ &\quad + C\|g\|_\beta \frac{1}{\varepsilon}\int_s^t\int_s^r \frac{|\bar{\sigma}_\varepsilon(r)-\bar{\sigma}_\varepsilon(\tau)|}{(r-\tau)^{\alpha+1}(t-r)^{1-\alpha-\beta}}d\tau dr\,. \end{aligned}$$

Use the same technique as for I_3 and I_2 to obtain

$$I_4 \leq C\|g\|_\beta\left(I_{41}+I_{42}+I_{43}+I_{44}\right)\,, \tag{4.34}$$

where

$$\begin{aligned} I_{41} &:= \int_s^t \frac{(t-r)^{\alpha+\beta-1}}{(r-s)^\alpha}\bigg|\int_0^1 \sigma_x(r,x^*(r)+\lambda y^\varepsilon(r),u^\varepsilon(r))d\lambda\cdot\eta^\varepsilon(r) \\ &\quad + \int_0^1 \left(\sigma_x(r,x^*(r)+\lambda y^\varepsilon(r),u^\varepsilon(r))-\sigma^*_x(r)\right)d\lambda\cdot y(r) \\ &\quad + \int_0^1 \left(\sigma_u(r,x^*(r),u^*(r)+\lambda\varepsilon\bar{u}(r))-\sigma^*_u(r)\right)d\lambda\cdot\bar{u}(r)\bigg|\,dr \\ I_{42} &= \int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\bigg|\int_0^1 \sigma_x(r,x^*(r)+\lambda y^\varepsilon(r),u^\varepsilon(r))d\lambda\cdot\eta^\varepsilon(r) \\ &\quad - \int_0^1 \sigma_x(\tau,x^*(\tau)+\lambda y^\varepsilon(\tau),u^\varepsilon(\tau))d\lambda\cdot\eta^\varepsilon(\tau)\bigg|\,d\tau dr \\ I_{43} &:= \int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\bigg|\int_0^1 \left(\sigma_x(r,x^*(r)+\lambda y^\varepsilon(r),u^\varepsilon(r))-\sigma^*_x(r)\right)d\lambda\cdot y^\varepsilon(r) \\ &\quad - \int_0^1 \left(\sigma_x(\tau,x^*(\tau)+\lambda y^\varepsilon(\tau),u^\varepsilon(\tau))-\sigma^*_x(\tau)\right)d\lambda\cdot y^\varepsilon(\tau)\bigg|\,d\tau dr \\ I_{44} &:= \int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\bigg|\int_0^1 \left(\sigma_u(r,x^*(r),u^*(r)+\lambda\varepsilon\bar{u}(r))-\sigma^*_u(r)\right)d\lambda\cdot(\varepsilon\bar{u}(r)) \\ &\quad - \int_0^1 \left(\sigma_u(\tau,x^*(\tau),u^*(\tau)+\lambda\varepsilon\bar{u}(\tau))-\sigma^*_u(\tau)\right)d\lambda\cdot(\varepsilon\bar{u}(\tau))\bigg|\,d\tau dr\,. \end{aligned}$$

We shall use the boundedness of the first derivatives of σ. Since the second derivatives of σ with respect to x and u are bounded the first derivatives of σ are Lipschitzian. Thus we have

$$\begin{aligned} I_{41} &\leq \int_s^t \frac{(t-r)^{\alpha+\beta-1}}{(r-s)^\alpha}|\eta^\varepsilon(r)|dr \\ &\quad + \int_s^t \frac{(t-r)^{\alpha+\beta-1}}{(r-s)^\alpha}\left[|y^\varepsilon(r)|+|\varepsilon\bar{u}(r)|\right]\left[|y(r)|+\bar{u}(r)\right]dr\,. \end{aligned}$$

Since $\sup_{0\le t\le T}|y^\varepsilon(t)|\le C\varepsilon$, and $\bar u(r)$ and $y(r)$ are bounded, we see

$$I_{41}\le C(t-s)^\beta\left[\|\eta^\varepsilon\|_{s,t,\infty}+\varepsilon\right]. \tag{4.35}$$

In a similar way we have

$$\begin{aligned} I_{42} \quad\le\quad & \int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}|\eta^\varepsilon(r)-\eta^\varepsilon(\tau)|d\tau dr \\ & +\|g\|_\beta\int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\left((r-\tau)^\gamma+|x^*(r)-x^*(\tau)|+\right. \\ & \left.+|x^\varepsilon(r)-x^\varepsilon(\tau)|+|u^*(r)-u^*(\tau)|+|u^\varepsilon(r)-u^\varepsilon(\tau)|\right)|\eta^\varepsilon(\tau)|\,d\tau dr\,. \end{aligned}$$

By the Hölder continuity of x^* and u^* and unform Hölder continuity of x^ε and u^ε, which are used to assure the integrability, we obtain (similar to $\tilde I_2$)

$$I_{42} \quad\le\quad (t-s)^{\beta+1-\alpha}\|\eta^\varepsilon\|_{1-\alpha,s,t}+(t-s)^{\beta+\nu}\|\eta^\varepsilon\|_{s,t,\infty}\,. \tag{4.36}$$

I_{43} and I_{44} are more complex and can be dealt with in the same way. We shall consider I_{43}. First we have

$$I_{43} \quad\le\quad I_{431}+I_{432}\,, \tag{4.37}$$

where

$$\begin{aligned} I_{431} \quad:=\quad & \int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\int_0^1 |\sigma_x(r,x^*(r)+\lambda y^\varepsilon(r),u^\varepsilon(r))-\sigma_x^*(r)|\,d\lambda \\ & \cdot|y^\varepsilon(r)-y^\varepsilon(\tau)|\,d\tau dr \\ I_{432} \quad:=\quad & \int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}\int_0^1 \Big|\sigma_x(r,x^*(r)+\lambda y^\varepsilon(r),u^\varepsilon(r))-\sigma_x^*(r) \\ & -\sigma_x(\tau,x^*(\tau)+\lambda y^\varepsilon(\tau),u^\varepsilon(\tau))+\sigma_x^*(\tau))\Big|d\lambda\cdot|y^\varepsilon(\tau)|\,d\tau dr\,. \end{aligned}$$

By the Hölder continuity of y^ε and Lipschitz continuity of the first derivatives of σ, we have

$$I_{431} \quad\le\quad \int_s^t\int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}}(r-\tau)^{1-\alpha}\left[|y^\varepsilon(r)|+\varepsilon|\bar u(r)|\right]dr\le C(t-s)^{\beta+1-\alpha}\varepsilon\,. \tag{4.38}$$

To deal with I_{432} we denote

$$x^\varepsilon_{\lambda,\eta}(r)=x^*(r)+\lambda\eta y^\varepsilon(r)\,,\quad u^\varepsilon_\eta(r)=u^*(r)+\eta\bar u(r)\,.$$

Then

$$\begin{aligned} & \sigma_x(r,x^*(r)+\lambda y^\varepsilon(r),u^*(r)+\varepsilon\bar u(r))-\sigma_x^*(r) \\ = \quad & \int_0^1\left[\lambda\sigma_{xx}(r,x^\varepsilon_{\lambda,\eta}(r),u^\varepsilon_\eta(r))y^\varepsilon(r)+\varepsilon\sigma_{xu}(r,x^\varepsilon_{\lambda,\eta}(r),u^\varepsilon_\eta(r))\bar u(r)\right]d\eta\,. \end{aligned}$$

Hence,

$$
\begin{aligned}
I_{432} \quad \leq \quad & \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} \int_0^1 \int_0^1 \bigg[\left|\sigma_{xx}(r, x^{\varepsilon}_{\lambda,\eta}(r), u^{\varepsilon}_{\eta}(r)) y^{\varepsilon}(r) - \sigma_{xx}(\tau, x^{\varepsilon}_{\lambda,\eta}(\tau), u^{\varepsilon}_{\eta}(\tau)) y^{\varepsilon}(\tau)\right| \\
& + \varepsilon \left|\sigma_{xu}(r, x^{\varepsilon}_{\lambda,\eta}(r), u^{\varepsilon}_{\eta}(r)) \bar{u}(r) - \sigma_{xu}(\tau, x^{\varepsilon}_{\lambda,\eta}(\tau), u^{\varepsilon}_{\eta}(\tau)) \bar{u}(\tau)\right| \bigg] d\lambda d\eta \cdot |y^{\varepsilon}(\tau)| \, d\tau dr
\end{aligned}
$$

By the boundedness and the Hölder continuity of the second derivatives of σ we have

$$
\begin{aligned}
I_{432} \quad \leq \quad & C \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} |y^{\varepsilon}(r) - y^{\varepsilon}(\tau)| \, |y^{\varepsilon}(\tau)| \, dr \\
& + C\varepsilon \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} |\bar{u}(r) - \bar{u}(\tau)| \, |y^{\varepsilon}(\tau)| \, dr \\
& + C \int_s^t \int_s^r \frac{(t-r)^{\alpha+\beta-1}}{(r-\tau)^{\alpha+1}} \bigg[(r-\tau)^{\gamma} + |x^*(r) - x^*(\tau)| + |y^{\varepsilon}(r) - y^{\varepsilon}(\tau)| \\
& + \varepsilon |u^*(r) - u^*(\tau)| + |u^{\varepsilon}(r) - u^{\varepsilon}(\tau)| \bigg] \cdot \bigg[|y^{\varepsilon}(\tau)| + \varepsilon \, |\bar{u}(\tau)| \bigg] |y^{\varepsilon}(\tau)| \, d\tau dr \, .
\end{aligned}
$$

By equation (4.30), we have

$$
I_{432} \leq C(t-s)^{\beta} \varepsilon \, . \tag{4.39}
$$

Combination of (4.37)-(4.39) yields

$$
I_{43} \leq C(t-s)^{\beta} \varepsilon \, . \tag{4.40}
$$

In similar way, we have

$$
I_{44} \leq C(t-s)^{\beta} \varepsilon \, . \tag{4.41}
$$

The inequalities (4.34), (4.35), (4.36), (4.40), and (4.41) yield

$$
I_4 \leq C \|g\|_{\beta} \left[(t-s)^{\beta+1-\alpha} \|\eta^{\varepsilon}\|_{1-\alpha,s,t} + (t-s)^{\beta} \|\eta^{\varepsilon}\|_{s,t,\infty} + \varepsilon (t-s)^{\beta} \right] \, . \tag{4.42}
$$

Step 6. From (4.32), (4.33) and (4.42), we have

$$
|\eta^{\varepsilon}(t) - \eta^{\varepsilon}(s)| \leq C \|g\|_{\beta} \left[(t-s)^{\beta+1-\alpha} \|\eta^{\varepsilon}\|_{1-\alpha,s,t} + (t-s)^{\beta} \|\eta^{\varepsilon}\|_{s,t,\infty} + \varepsilon (t-s)^{\beta} \right] \, .
$$

This implies

$$
\|\eta^{\varepsilon}\|_{1-\alpha,s,t} \leq C \|g\|_{\beta} \left[(t-s)^{\beta} \|\eta^{\varepsilon}\|_{1-\alpha,s,t} + (t-s)^{\beta+\alpha-1} \|\eta^{\varepsilon}\|_{s,t,\infty} + \varepsilon (t-s)^{\beta+\alpha-1} \right] \, .
$$

Now we can follow the same argument as in Step 3 to complete the proof of the theorem.

Step 7. The L^p convergence is from the Fernique Theorem. ■

As we mentioned earlier Lemma 4.4 implies Proposition 4.1 easily. This completes the proof for Proposition 4.1. ■

To derive the maximum principle, we need to obtain an explicit solution to the equation (4.6). Let us consider the following linear matrix-valued stochastic differential equations.

$$\begin{cases} d\Phi(t) &=& b_x^*(t)\Phi(t)dt + \sum_{j=1}^m \sigma_x^{j,*}(t)\Phi(t)d^\circ B_j^H(t), \\ \Phi(0) &=& I, \end{cases} \tag{4.43}$$

From the basic stochastic calculus for fractional Brownian motions (see e.g.[22]), it is clear that this equation has a unique solution, denoted by $\Phi(t)$. It is easy to verify that $\Phi^{-1}(t)$ exists and is the unique solution of the following stochastic differential equations:

$$\begin{cases} d\Phi^{-1}(t) &=& -\Phi^{-1}(t)b_x^*(t)dt - \sum_{j=1}^m \Phi^{-1}(t)\sigma_x^{j,*}(t)d^\circ B_j^H(t), \\ \Phi^{-1}(0) &=& I\,. \end{cases} \tag{4.44}$$

Again from the Itô's formula we can obtain the solution of the equation (4.6) by using $\Phi(t)$ and $\Phi^{-1}(t)$.

Lemma 4.5 *Let $\Phi(t)$ and $\Phi^{-1}(t)$ be defined by (4.43) and (4.44). Then the solution to (4.6) is given by*

$$\begin{aligned} y(t) &= \Phi(t)\int_0^t \Phi^{-1}(s)b_u^*(s)\bar{u}(s)ds \\ &\quad + \sum_{j=1}^m \Phi(t)\int_0^t \Phi^{-1}(s)\sigma_u^{j,*}(s)\bar{u}(s)dB_j^H(s)\,, \end{aligned} \tag{4.45}$$

5 Maximum principle for stochastic control of system driven by fractional Brownian motion

Recall that we defined the space of admissible controls in Section 4,

$$\begin{aligned} U[0,T] \;:=\; \Big\{ & u | u : [0,T]\times\Omega \to \mathbb{R}^d, u \text{ is } \mathcal{G}\text{-adapted, } u \in C^\mu[0,T] \text{ for some } \mu > 1-H\,, \\ & \text{there exist constant } C>0,\ c>0, \text{ and } \beta < H \\ & \quad \text{such that } \ \|u\|_{\mu,0,T} \le Ce^{c\sum_{j=1}^m \|B_j^H\|_{\beta,0,T}} \\ & \quad \text{and } \ \int_0^T\int_0^T \mathbb{E}\left(|D_s^H u(t)|^2\right) dsdt < \infty \Big\}. \end{aligned} \tag{5.1}$$

It is clear that $U[0,T]$ is a linear space.

Let $b : [0,T]\times\mathbb{R}^n\times\mathbb{R}^d \to \mathbb{R}^n$, $\sigma : [0,T]\times\mathbb{R}^n\times\mathbb{R}^d \to \mathbb{R}^{n\times m}$, be some given continuous functions satisfying the assumptions (H4) and (H5) given in Section

4. Consider the following controlled system of stochastic differential equations driven by fractional Brownian motions:

$$\begin{cases} dx(t) &= b(t,x(t),u(t))dt + \sum_{j=1}^{m} \sigma^j(t,x(t),u(t))d^\circ B_j^H(t), \\ x(0) &= x_0 \,. \end{cases} \tag{5.2}$$

Here the integral with respect to fractional Brownian motion is the Stratonovich integral. Some properties of this controlled system are given in Section 4.

Let $l : [0,T]\times\mathbb{R}^n\times\mathbb{R}^d \ \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}$ be some given functions satisfying the following conditions.

(**H3**) l and h are continuously differentiable with bounded derivatives.

The cost functional we use in this paper is given by

$$J(u(\cdot)) = \mathbb{E}\left\{\int_0^T l(t,x(t),u(t))dt + h(x(T))\right\}. \tag{5.3}$$

Assume $\gamma > 1-H$. Let $\alpha \in (1-H, \alpha_0)$ and $\rho \geq 1/\alpha$, where

$$\alpha_0 = \min\left\{\frac{1}{2}, \gamma\right\}.$$

From the conditions (H4)-(H5) the controlled stochastic differential equation (5.2) has a unique solution (see e.g. [17], [28]). Moreover for P-almost all $\omega \in \Omega$, $X(\omega,\cdot) \in C^{1-\alpha}(0,T,\mathcal{R}^d)$. So assume that $|x(r)-x(\tau)| \leq c_0|r-\tau|^{1-\alpha}$ in probability. The solution x_t^u of above equation depends on u. But to simplify the notation we often omit its explicit dependence on u and write $x_t = x_t^u$.

Now our optimal control problem can be stated as to minimize the cost functional over $U[0,T]$. That is to find optimal control $u^*(\cdot) \in U[0,T]$ such that

$$J(u^*(\cdot)) = \inf_{u(\cdot)\in U[0,T]} J(u(\cdot)). \tag{5.4}$$

Let $u^*(t)$ be an optimal control and $x^*(t)$ be the corresponding solution of equation (5.2). $(x^*(\cdot), u^*(\cdot))$ is called an optimal pair.

We will find a necessary condition that the optimal control $u^*(\cdot)$ must satisfy, which is also called the maximum principle.

5.1 Stochastic control with partial information

Since $u^*(\cdot)$ is an optimal control and since the space $U[0,T]$ of admissible controls is linear, we have that $u^*(\cdot)$ is a critical point of nonlinear functional $J(u(\cdot))$, $u \in U[0,T]$. This means that

$$\frac{d}{d\varepsilon} J(u^*(\cdot)+\varepsilon\bar{u}(\cdot))\Big|_{\varepsilon=0} = 0\,. \tag{5.5}$$

We will follow the same idea as in Section 3 for the Brownian motion case to deduce a necessary condition from the above equation (5.5). As in Section 4, we denote $y(r) = \lim_{\varepsilon\to 0} \frac{x^\varepsilon(r) - x^*(r)}{\varepsilon}$. By Lemma (4.4) and Dominated Convergence Theorem, we have

$$\frac{d}{d\varepsilon} J(u^*(\cdot) + \varepsilon \bar{u}(\cdot))\Big|_{\varepsilon=0} = \mathbb{E}\int_0^T \left(l_x^{*T}(t) y(t) + l_u^{*T}(t)\bar{u}(t)\right) dt + \mathbb{E}\left[h_x^T(x^*(T)) y(T)\right] .$$

Substituting (4.21) into (5.5) we obtain

$$\begin{aligned}
& \frac{d}{d\varepsilon} J(u^*(\cdot) + \varepsilon \bar{u}(\cdot))\Big|_{\varepsilon=0} = \mathbb{E}\int_0^T \left(l_x^{*T}(t)\Phi(t)\int_0^t \Phi^{-1}(s) b_u^*(s)\bar{u}(s) ds + l_u^{*T}(t)\bar{u}(t)\right) dt \\
& + \mathbb{E}\left\{ h_x^T(x^*(T))\Phi(T)\int_0^T \Phi^{-1}(s) b_u^*(s)\bar{u}(s) ds\right\} \\
& + \sum_{j=1}^m \mathbb{E}\int_0^T \left(l_x^{*T}(t)\Phi(t)\int_0^t \Phi^{-1}(s)\sigma_u^{j,*}(s)\bar{u}(s) d^\circ B_j^H(s)\right) dt \\
& + \sum_{j=1}^m \mathbb{E}\left\{ h_x^T(x^*(T))\Phi(T)\int_0^T \Phi^{-1}(s)\sigma_u^{j,*}(s)\bar{u}(s) d^\circ B_j^H(s)\right\} \\
= \; & \mathbb{E}\int_0^T \left(\int_s^T \left(l_x^{*T}(t)\Phi(t)\right) dt \Phi^{-1}(s) b_u^*(s)\bar{u}(s)\right) ds + \mathbb{E}\int_0^T l_u^{*T}(s)\bar{u}(s) ds \\
& + \mathbb{E}\int_0^T h_x^T(x^*(T))\Phi(T)\Phi^{-1}(s) b_u^*(s)\bar{u}(s) ds \\
& + \sum_{j=1}^m \int_0^T \left(\mathbb{E} l_x^{*T}(t)\Phi(t)\int_0^t \left(\Phi^{-1}(s)\sigma_u^{j,*}(s)\bar{u}(s) d^\circ B_j^H(s)\right) ds\right) dt \\
& + \sum_{j=1}^m \mathbb{E}\left\{ h_x^T(x^*(T))\Phi(T)\int_0^T \Phi^{-1}(s)\sigma_u^{j,*}(s)\bar{u}(s) d^\circ B_j^H(s)\right\} .
\end{aligned}$$

The expectation of the above last two terms can be computed by the formula (2.19). Thus, we have

$$
\begin{aligned}
& \frac{d}{d\varepsilon} J(u^*(\cdot)+\varepsilon \bar{u}(\cdot))\Big|_{\varepsilon=0} \\
= \quad & \mathbb{E}\int_0^T \left(\int_s^T {l_x^*}^\top(t)\Phi(t)dt\right)\Phi^{-1}(s)b_u^*(s)\bar{u}(s)ds + \mathbb{E}\int_0^T {l_u^*}^\top(s)\bar{u}(s)ds \\
& +\mathbb{E}\int_0^T h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)b_u^*(s)\bar{u}(s)ds \\
& \quad +\sum_{j=1}^m \mathbb{E}\int_0^T \left(\int_s^T \mathbb{D}_s^j\left\{{l_x^*}^\top(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right\}\bar{u}(s)dt\right)ds \\
& +\sum_{j=1}^m \mathbb{E}\int_0^T\int_s^T {l_x^*}^\top(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)\mathbb{D}_s^j\left(\bar{u}(s)\right)dtds \\
& \quad +\sum_{j=1}^m \mathbb{E}\int_0^T \mathbb{D}_s^j\left(h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right)\bar{u}(s)ds \\
& +\sum_{j=1}^m \mathbb{E}\int_0^T h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\mathbb{D}_s^j\left(\bar{u}(s)\right)ds.
\end{aligned}
$$

Since the equation (5.5) holds true for all adapted process u in $U[0,T]$, we can choose especially $\bar{u}(s) = \mathbf{1}_{[a,b]}\tilde{u}$, where $0 \le a \le b \le T$ and $\tilde{u}$ which is $\mathcal{G}_a$ measurable. Then from (5.5) and the above computation it follows

$$
\begin{aligned}
& \mathbb{E}\int_a^b \left(\int_s^T {l_x^*}^\top(t)\Phi(t)dt\right)\Phi^{-1}(s)b_u^*(s)\bar{u}ds + \mathbb{E}\int_a^b {l_u^*}^\top(s)\bar{u}ds \\
& +\mathbb{E}\int_a^b h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)b_u^*(s)\bar{u}ds \\
& \quad +\sum_{j=1}^m \mathbb{E}\int_a^b \left(\int_s^T \mathbb{D}_s^j\left\{{l_x^*}^\top(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right\}\bar{u}dt\right)ds \\
& +\sum_{j=1}^m \mathbb{E}\int_a^b\int_s^T {l_x^*}^\top(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)\mathbb{D}_s^j\left(\bar{u}\right)dtds \\
& \quad +\sum_{j=1}^m \mathbb{E}\int_a^b \mathbb{D}_s^j\left(h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right)\bar{u}ds \\
& +\sum_{j=1}^m \mathbb{E}\int_a^b h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\mathbb{D}_s^j\left(\bar{u}\right)ds = 0.
\end{aligned}
$$

We can use the formula (2.17) to compute the above two terms involving $\mathbb{D}_s^j \bar{u}$. we have then

$$\mathbb{E}\Bigg[\Bigg\{\int_a^b \left(\int_s^T l_x^{*\top}(t)\Phi(t)dt\right)\Phi^{-1}(s)b_u^*(s)ds + \int_a^b l_u^{*\top}(s)ds$$

$$+\int_a^b h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)b_u^*(s)ds$$

$$+\sum_{j=1}^m \int_a^b \left(\int_s^T \mathbb{D}_s^j\left\{l_x^{*\top}(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right\}dt\right)ds$$

$$+\sum_{j=1}^m \int_a^b \left[\int_s^T l_x^{*\top}(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)dt\right]dB_j^H(s)$$

$$+\sum_{j=1}^m \int_a^b \mathbb{D}_s^j\left(h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right)\bar{u}ds$$

$$+\sum_{j=1}^m \int_a^b h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)dB_j^H(s)\Bigg\}\bar{u}\Bigg] = 0.$$

Since the $\mathcal{G}_a$ measurable $\bar{u}$ is arbitrary, we have

Theorem 5.1 *Let u^* be the optimal admissible control. Then u^* satisfies the following*

$$\mathbb{E}\Bigg[\Bigg\{\int_a^b \left[\int_s^T l_x^{*\top}(t)\Phi(t)dt\Phi^{-1}(s)b_u^*(s) + l_u^{*\top}(s) + h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)b_u^*(s)\right]ds$$

$$+\sum_{j=1}^m \int_a^b \Bigg[\int_s^T \mathbb{D}_s^j\left\{l_x^{*\top}(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right\}dt$$

$$+\mathbb{D}_s^j\left\{h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\right\}\Bigg]ds + \sum_{j=1}^m \int_a^b \Bigg[\int_s^T l_x^{*\top}(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)dt$$

$$+h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\Bigg]dB_j^H(s)\Bigg\}\Bigg|\mathcal{G}_a\Bigg] = 0\,, \quad \forall\ \ 0\le a\le b\le T\,. \tag{5.6}$$

Remark 5.2 *Using pathwise integral, we can write the above equation as*

$$\mathbb{E}\Bigg[\Bigg\{\int_a^b \left[\int_s^T l_x^{*\top}(t)\Phi(t)dt\Phi^{-1}(s)b_u^*(s) + l_u^{*\top}(s) + h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)b_u^*(s)\right]ds$$

$$\sum_{j=1}^m \int_a^b \Bigg[\int_s^T l_x^{*\top}(t)\Phi(t)\Phi^{-1}(s)\sigma_u^{j,*}(s)dt$$

$$+h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\Bigg]d^\circ B_j^H(s)\Bigg\}\Bigg|\mathcal{G}_a\Bigg] = 0\,, \quad \forall\ \ 0\le a\le b\le T\,. \tag{5.7}$$

5.2 Stochastic control with complete information

Now we assume that the filtration $\mathcal{G} = \mathcal{F}$ and note that it is also the filtration generated by the (background) defining Brownian motions W. Namely,

$$\mathcal{F}_t = \sigma\left(B_1^H(s), \cdots, B_m^H(s),\ 0 \le s \le t\right) = \sigma\left(W_1(s), \cdots, W_m(s),\ 0 \le s \le t\right).$$

We shall simplify the equation (5.6). Denote

$$\begin{aligned}
F(T,s) &= \left(\int_s^T l_x^{*\top}(t)\Phi(t)dt\right)\Phi^{-1}(s)b_u^*(s) + l_u^{*\top}(s) + h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)b_u^*(s)\,,\\
G_j(T,s) &= \left(\int_s^T l_x^{*\top}(t)\Phi(t)dt\right)\Phi^{-1}(s)\sigma_u^{j,*}(s) + h_x^\top(x^*(T))\Phi(T)\Phi^{-1}(s)\sigma_u^{j,*}(s)\,,\\
\tilde{F}(T,s) &= F(T,s) + \sum_{j=1}^m \mathbb{D}_s^j G_j(T,s)\,.
\end{aligned}$$

Then the equation (5.6) can be written as

$$\mathbb{E}\left\{\int_a^b \tilde{F}(T,s)ds + \sum_{j=1}^m \int_a^b G_j(T,s)dB_j^H(s)\Big|\mathcal{F}_a\right\} = 0\,. \tag{5.1}$$

We shall simplify the above equation (5.1) in the case of complete information. To this end we need some lemmas.

Lemma 5.3 *Let $1/2 < H < 1$, and $\varepsilon > 0$. Denote*

$$\rho(\varepsilon) = \int_0^a \int_0^a (t-s)^{2H-2} s^{\frac{1}{2}-H}(a-s+\varepsilon)^{-H-\frac{1}{2}} t^{\frac{1}{2}-H}(a-t+\varepsilon)^{-H-\frac{1}{2}} dsdt.$$

For any $\delta \in (0, 2-2H)$, there exist $C_\delta > 0$, depending on H, a and δ but independent of ε, such that

$$\rho(\varepsilon) \ge C_\delta \varepsilon^{1-2H-\delta}\,. \tag{5.2}$$

Proof Without loss of generality, we prove the case when $a = 1$.
We choose some $\delta \in (0, 2-2H)$, and we have $|t-s|^{2H-2}(1-s+\varepsilon)^\delta = |t-$

$s|^{2H-2+\delta}(\frac{1-s+\varepsilon}{t-s})^{\delta} \geq 1$, when $0 \leq s \leq t \leq 1$.

$$
\begin{aligned}
&\int_0^1 \int_0^1 |t-s|^{2H-2} s^{\frac{1}{2}-H} t^{\frac{1}{2}-H} (1-s+\varepsilon)^{-H-\frac{1}{2}} (1-t+\varepsilon)^{-H-\frac{1}{2}} dsdt \\
=&2 \int_0^1 \int_0^t |t-s|^{2H-2} s^{\frac{1}{2}-H} t^{\frac{1}{2}-H} (1-s+\varepsilon)^{-H-\frac{1}{2}} (1-t+\varepsilon)^{-H-\frac{1}{2}} dsdt \\
\geq&2 \int_0^1 \int_0^t |t-s|^{2H-2} (1-s+\varepsilon)^{-H-\frac{1}{2}} (1-t+\varepsilon)^{-H-\frac{1}{2}} dsdt \\
\geq&2 \int_0^1 \int_0^t (1-s+\varepsilon)^{-H-\frac{1}{2}-\delta} (1-t+\varepsilon)^{-H-\frac{1}{2}} dsdt \\
=&C \int_0^1 \left[(1-t+\varepsilon)^{\frac{1}{2}-H-\delta} - (1+\varepsilon)^{\frac{1}{2}-H-\delta} \right] (1-t+\varepsilon)^{-H-\frac{1}{2}} dt \\
=&C \left(\frac{1}{2H+\delta-1} [\varepsilon^{1-2H-\delta} - (1+\varepsilon)^{1-2H-\delta}] - \frac{1}{H-\frac{1}{2}} [\varepsilon^{-H+\frac{1}{2}} - (1+\varepsilon)^{-H+\frac{1}{2}}] (1+\varepsilon)^{\frac{1}{2}-H-\delta} \right) \\
\sim&\varepsilon^{1-2H-\delta} \text{ as } \varepsilon \to 0.
\end{aligned}
$$

■

Lemma 5.4 *Let X_t be a Gaussion random variable with 0 mean and variance $f^2(t)$. If $\lim_{t\to 0} f(t) = \infty$, then $\lim_{t\to 0} |X_t| = \infty$ in probability.*

Proof We have

$$
\begin{aligned}
E(e^{-|X_t|}) &= \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-f(t)x} e^{-\frac{x^2}{2}} dx \\
&\leq \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-f(t)x} dx = \frac{2}{\sqrt{2\pi}} \frac{1}{f(t)} .
\end{aligned}
$$

This implies that $\lim_{t\to 0} E(e^{-|X_t|}) = 0$ and hence $e^{-|X_t|}$ goes to 0 in L^1. Consequently, we have $\lim_{t\to 0} |X_t| = \infty$ in probability as $t \to 0$. ■

Lemma 5.5 *Let g be continuous on $[0,T]$ and let f_j, $j = 1, 2, \cdots, m$ be Hölder continuous on $[0,T]$ of order ρ with $\rho > 1 - H$. Assume that the Malliavin derivative $\mathbb{D}_s^j f_j(t)$ is continuous in $s \in [0,T]$ for all $t \in [0,T]$ and assume that $\sup_{0\leq s\leq T} \mathbb{D}_s^j \mathbb{E}\left[f_j(t)|\mathcal{F}_t\right] < \infty$ almost surely for all $t \in [0,T]$. If*

$$
\mathbb{E}\left\{ \int_a^b g(s)ds + \sum_{j=1}^m \int_a^b f_j(s) d^\circ B_j^H(s) \Big| \mathcal{F}_a \right\} = 0 , \quad \forall \;\; 0 \leq a \leq b \leq T , \tag{5.3}
$$

then for all $j = 1, 2, \cdots, m$,

$$
\mathbb{E}\left[f_j(a)|\mathcal{F}_a\right] = 0 , \quad \forall \;\; 0 < a \leq T . \tag{5.4}
$$

Proof

Step 1. Let $\varepsilon > 0$ and $0 < a < a + \varepsilon < T$. The equation (5.3) can be written as

$$\mathbb{E}\left\{ \int_a^{a+\varepsilon} g(s)ds + \sum_{j=1}^m \int_a^{a+\varepsilon} [f_j(s) - f_j(a)]\, d^\circ B_j^H(s) \Big| \mathcal{F}_a \right\}$$
$$+ \sum_{j=1}^m \mathbb{E}\left\{ f_j(a) \left[B_j^H(a+\varepsilon) - B_j^H(a) \right] \Big| \mathcal{F}_a \right\} = 0 \qquad (5.5)$$

Let $0 < \beta < H$ such that $\beta + \rho > 1$. Then from (2.5), it follows that

$$\left| \int_a^{a+\varepsilon} [f_j(s) - f_j(a)]\, d^\circ B_j^H(s) \right| \;\; \le \;\; \|B_j^H\|_{\beta,0,T} \varepsilon^{\beta+\rho} .$$

Dividing (5.5) by $\varepsilon^{H'}$ with $0 < H' < 1$, then we have

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^{H'}} \sum_{j=1}^m \mathbb{E}\left\{ f_j(a) \left[B_j^H(a+\varepsilon) - B_j^H(a) \right] \Big| \mathcal{F}_a \right\} = 0 . \qquad (5.6)$$

It is maybe possible to compute the above expectation in an easy way. But instead of developing a formula for above conditional expectation we shall use the results from [14]. First, we have

$$\int_a^{a+\varepsilon} f_j(a) dB_j^H(s) = f_j(a) \left[B_j^H(a+\varepsilon) - B_j^H(a) \right] - \int_a^{a+\varepsilon} \mathbb{D}_s^j f_j(a) ds$$

By the continuity of $\mathbb{D}_s f_j(a)$, (5.6) implies

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^{H'}} \sum_{j=1}^m \mathbb{E}\left\{ \int_a^{a+\varepsilon} f_j(a) dB_j^H(s) \Big| \mathcal{F}_a \right\} = 0 , \quad \forall a \in [0, T] . \qquad (5.7)$$

Step 2. Let us recall some notations in [14] (see Equations (5.21), (5.34), and (9.22) of [14]).

$$\mathbf{\Gamma}^*_{H,T} f(t) = (H - \frac{1}{2}) \kappa_H t^{\frac{1}{2}-H} \int_t^T \xi^{H-\frac{1}{2}} (\xi - t)^{H-\frac{3}{2}} f(\xi) d\xi ,$$
$$\mathbb{B}^*_{H,\tau} g(t) = -\frac{2H\kappa_1}{\kappa_H} t^{\frac{1}{2}-H} \frac{d}{dt} \int_t^\tau (\eta - t)^{\frac{1}{2}-H} \eta^{H-\frac{1}{2}} g(\eta) d\eta ,$$

and

$$\mathbb{P}_{H,\tau}(t) f(t) = \mathbb{B}^*_{H,\tau} \mathbf{\Gamma}^*_{H,T} f(t) .$$

Then from the equation (9.21) of [14], we have

$$
\begin{aligned}
&\mathbb{E}\left[\int_a^{a+\varepsilon} f_j(a)dB_j^H(s)\Big|\,\mathcal{F}_a\right] \\
= &\int_0^a \mathbb{P}_{H,a}(s)\mathbb{E}\left[\mathbf{1}_{[a,a+\varepsilon]}(s)f_j(a)\big|\,\mathcal{F}_a\right]dB_j^H(s) \\
= &\int_0^a \left\{Cs^{\frac{1}{2}-H}\frac{d}{ds}\int_s^a \left[(\eta-s)^{\frac{1}{2}-H}\eta^{H-\frac{1}{2}}(H-\frac{1}{2})\kappa_H\eta^{\frac{1}{2}-H}\right.\right. \\
&\left.\left.\cdot\int_\eta^T \left(\xi^{H-\frac{1}{2}}(\xi-\eta)^{H-\frac{3}{2}}\mathbb{E}\left[\mathbf{1}_{[a,a+\varepsilon]}(\xi)f_j(a)|\mathcal{F}_a\right]\right)d\xi\right]d\eta\right\}dB_j^H(s) \\
= &\ C\int_0^a \zeta(a,\varepsilon,s)\mathbb{E}\left[f_j(a)|\mathcal{F}_a\right]dB_j^H(s) \\
= &\ C\mathbb{E}\left[f_j(a)|\mathcal{F}_a\right]\int_0^a \zeta(a,\varepsilon,s)dB_j^H(s) - C\int_0^a \zeta(a,\varepsilon,s)\mathbb{D}_s^j\left(\mathbb{E}(f_j(a)|\mathcal{F}_a]\right)ds\,,
\end{aligned}
\tag{5.8}
$$

where

$$
\zeta(a,\varepsilon,s) = s^{\frac{1}{2}-H}\frac{d}{ds}\int_s^a (\eta-s)^{\frac{1}{2}-H}\left[\int_a^{a+\varepsilon}\xi^{H-\frac{1}{2}}(\xi-\eta)^{H-\frac{3}{2}}d\xi\right]d\eta\,.
$$

Step 3. This function ζ can be calculated as follows.

$$
\begin{aligned}
\zeta(a,\varepsilon,s) &= s^{\frac{1}{2}-H}\frac{d}{ds}\int_a^{a+\varepsilon}\int_s^a (\eta-s)^{\frac{1}{2}-H}(\xi-\eta)^{H-\frac{3}{2}}\xi^{H-\frac{1}{2}}d\eta d\xi \\
&= s^{\frac{1}{2}-H}\frac{d}{ds}\int_a^{a+\varepsilon}\left(\int_s^\xi (\eta-s)^{\frac{1}{2}-H}(\xi-\eta)^{H-\frac{3}{2}}\ d\eta\right. \\
&\qquad \left.-\int_a^\xi (\eta-s)^{\frac{1}{2}-H}(\xi-\eta)^{H-\frac{3}{2}}\ d\eta\right)\xi^{H-\frac{1}{2}}d\xi \\
&= -s^{\frac{1}{2}-H}\frac{d}{ds}\int_a^{a+\varepsilon}\int_a^\xi (\eta-s)^{\frac{1}{2}-H}(\xi-\eta)^{H-\frac{3}{2}}d\eta\xi^{H-\frac{1}{2}}d\xi \\
&= s^{\frac{1}{2}-H}\int_a^{a+\varepsilon}\int_a^\xi (\frac{1}{2}-H)(\eta-s)^{-\frac{1}{2}-H}(\xi-\eta)^{H-\frac{3}{2}}d\eta\xi^{H-\frac{1}{2}}d\xi \\
&= \varepsilon^{H+\frac{1}{2}}s^{\frac{1}{2}-H}\int_0^1\int_0^1 (\eta'\xi'\varepsilon+a-s)^{-\frac{1}{2}-H}(1-\eta')^{H-\frac{3}{2}}\xi'^{H-\frac{1}{2}} \\
&\qquad (\xi'\varepsilon+a)^{H-\frac{1}{2}}d\eta'd\xi'\,.
\end{aligned}
$$

Choose $\gamma\in(0,1)$. Dividing (5.8) by ε^γ, we see

$$
\lim_{\varepsilon\to 0}\sum_{j=1}^m \mathbb{E}\left[f_j(a)|\mathcal{F}_a\right]\frac{1}{\varepsilon^\gamma}\int_0^a \zeta(a,\varepsilon,s)dB_j^H(s) = 0\,.
$$

But

$$\frac{1}{\varepsilon^\gamma}\int_0^a \zeta(a,\varepsilon,s)dB_j^H(s) = \varepsilon^{H+\frac12-\gamma}\int_0^1\int_0^1\left(\int_0^a s^{\frac12-H}(\eta\xi\varepsilon+a-s)^{-\frac12-H}dB_j(s)^H\right) \\ \times(1-\eta)^{H-\frac32}\xi^{H-\frac12}(\xi\varepsilon+a)^{H-\frac12}d\eta d\xi$$

and hence

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{\varepsilon^\gamma}\int_0^a \zeta(a,\varepsilon,s)dB_j^H(s)\right)^2 \\
=&C\varepsilon^{2H+1-2\gamma}\int_{[0,1]^4}\int_0^a\int_0^a |t-s|^{2H-2}s^{\frac12-H}t^{\frac12-H}(\eta_1\xi_1\varepsilon+a-s)^{-\frac12-H}(\eta_2\xi_2\varepsilon+a-t)^{-\frac12-H}dsdt \\
&\qquad\times\prod_{i=1}^2(1-\eta_i)^{H-\frac32}\xi_i^{H-\frac12}(\xi_i\varepsilon+a)^{H-\frac12}d\eta_1 d\xi_1 d\eta_2 d\xi_2 \\
\geq&C\varepsilon^{2H+1-2\gamma}\int_{[0,1]^4}\int_0^a\int_0^a |t-s|^{2H-2}s^{\frac12-H}t^{\frac12-H}(\varepsilon+a-s)^{-\frac12-H}(\varepsilon+a-t)^{-\frac12-H}dsdt \\
&\qquad\times\prod_{i=1}^2(1-\eta_i)^{H-\frac32}\xi_i^{H-\frac12}(\xi_i\varepsilon+a)^{H-\frac12}d\eta_1 d\xi_1 d\eta_2 d\xi_2 \\
=&C_\delta\varepsilon^{2H+1-\gamma+1-2H-\delta} \\
=&C_\delta\varepsilon^{2-2\gamma-\delta}
\end{aligned}$$

We may choose $\gamma\in(0,1)$ and $\delta\in(0,2-2H)$ such that $2-2\gamma-\delta<0$. From Lemma 5.4 we see that $\frac{1}{\varepsilon^\gamma}\int_0^a\zeta(a,\varepsilon,s)dB_j^H(s)$ converges to ∞ in probability as $\varepsilon\to 0$. This implies that $\mathbb{E}\left[f_j(a)|\mathcal{F}_a\right]=0$ a.s.. ■

Then we have also the following lemma.

Lemma 5.6 *Let g be continuous on $[0,T]$ and let f_j , $j=1,2,\cdots,m$ be Hölder continuous on $[0,T]$ of order ρ with $\rho>1-H$. Assume that the Malliavin derivative $\mathbb{D}_s^j f_j(t)$ is continuous in $s\in[0,T]$ for all $t\in[0,T]$ and assume that $\sup_{0\le s\le T}\mathbb{D}_s^j\mathbb{E}\left[f_j(t)|\mathcal{F}_t\right]<\infty$ almost surely for all $t\in[0,T]$. If*

$$\mathbb{E}\left\{\int_a^b g(s)ds+\sum_{j=1}^m\int_a^b f_j(s)dB_j^H(s)\Big|\mathcal{F}_a\right\}=0\,,\quad\forall\ \ 0\le a\le b\le T\,,\tag{5.9}$$

then for all $j=1,2,\cdots,m$,

$$\mathbb{E}\left[f_j(a)|\mathcal{F}_a\right]=0$$

and

$$\mathbb{E}\left[g(a)|\mathcal{F}_a\right]=0$$

for all $0<a\le T$.

Proof The first equality is obvious from the previous lemma. Now we prove the second one. For any $\varepsilon > 0$,

$$
\begin{aligned}
&\mathbb{E}\left[\int_a^{a+\varepsilon} f_j(s)dB_j^H(s)\bigg|\,\mathcal{F}_a\right]\\
=&\int_0^a \mathbb{P}_{H,a}(s)\mathbb{E}\left[\mathbf{1}_{[a,a+\varepsilon]}(s)f_j(a)\big|\,\mathcal{F}_a\right]dB_j^H(s)\\
&\cdot\int_\eta^T \left(\xi^{H-\frac{1}{2}}(\xi-\eta)^{H-\frac{3}{2}}\mathbb{E}\left[\mathbf{1}_{[a,a+\varepsilon]}(\xi)f_j(\xi)|\mathcal{F}_a\right]\right)d\xi\Bigg]\,d\eta\Bigg\}\,dB_j^H(s)\\
=&\int_0^a \mathbb{P}_{H,a}(s)\mathbb{E}\left[\mathbf{1}_{[a,a+\varepsilon]}(s)f_j(a)\big|\,\mathcal{F}_a\right]dB_j^H(s)\\
&\cdot\int_\eta^T \left(\xi^{H-\frac{1}{2}}(\xi-\eta)^{H-\frac{3}{2}}\mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{[a,a+\varepsilon]}(\xi)f_j(\xi)|\mathcal{F}_\xi|\mathcal{F}_a\right]\right]\right)d\xi\Bigg]\,d\eta\Bigg\}\,dB_j^H(s)\\
=&0,
\end{aligned}
$$

then we have

$$\mathbb{E}[\int_a^{a+\varepsilon} g(s)ds|\mathcal{F}_a] = 0, \forall\varepsilon > 0,$$

and hence

$$\mathbb{E}[g(a)|\mathcal{F}_a] = 0.$$

■

We apply Lemma 5.6 to the equation (5.1) and obtain

$$\mathbb{E}\left[G_j(T,t)\Big|\mathcal{F}_t\right] = 0\,, \tag{5.10}$$

and

$$\mathbb{E}\left[\tilde{F}(T,t)\Big|\mathcal{F}_t\right] = 0\,, \tag{5.11}$$

for all $0 < t \le T$. Denote

$$
\begin{aligned}
P(t) &= \left(\Phi^\top\right)^{-1}(t)\left[\int_t^T \Phi^\top(s)l_x^*(s)ds + \Phi^\top(T)h_x(x^*(T))\right]\\
p(t) &= \left(\Phi^\top\right)^{-1}(t)\mathbb{E}^{\mathcal{F}_t}\left[\int_t^T \Phi^\top(s)l_x^*(s)ds + \Phi^\top(T)h_x(x^*(T))\right].
\end{aligned}\tag{5.12}
$$

Then the equations (5.10) and (5.11) can be written as

$$\sum_{j=1}^m \sigma_u^{j,*\top}(t)p(t) = 0\,, \quad \forall\ 0 \le t \le T \tag{5.13}$$

and

$$b_u^{*\top}(t)p(t) + l_u^*(t) + \sum_{j=1}^{m} \mathbb{E}\left[\mathbb{D}_t^j G_j(T,t)\Big|\mathcal{F}_t\right] = 0\,, \quad \forall\ 0 \le t \le T\,. \tag{5.14}$$

Now we compute $\mathbb{E}\left[\mathbb{D}_t^j G_j(T,t)\Big|\mathcal{F}_t\right]$. Unlike in the classical case, which we have $\mathbb{E}\left[D_t^j G_j(T,t)\Big|\mathcal{F}_t\right] = D_t^j \mathbb{E}\left[G_j(T,t)\Big|\mathcal{F}_t\right]$, now we usually have $\mathbb{E}\left[\mathbb{D}_t^j G_j(T,t)\Big|\mathcal{F}_t\right] \neq \mathbb{D}_t^j \mathbb{E}\left[G_j(T,t)\Big|\mathcal{F}_t\right]$. We need to use (2.21). Let $c_{1,H}$ be a constant as defined in proposition 2.4 and $\phi_{1,H}(s,t) = c_{1,H} s^{\frac12 - H}|t-s|^{2H-2}$. Then

$$\begin{aligned}
\mathbb{E}\left[\mathbb{D}_t^j G_j(T,t)\Big|\mathcal{F}_t\right] &= \int_0^T \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_s^T (r-s)^{\frac12 - H} r^{H-\frac12}\mathbb{E}\left[D_r^j G_j(T,t)\Big|\mathcal{F}_t\right] dr\right) ds \\
&= I_1 + I_2\,,
\end{aligned} \tag{5.15}$$

where

$$\begin{aligned}
I_1 &= \int_0^t \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_s^t (r-s)^{\frac12 - H} r^{H-\frac12}\mathbb{E}\left[D_r^j G_j(T,t)\Big|\mathcal{F}_t\right] dr\right) ds \\
I_2 &= \int_0^T \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_{t\vee s}^T (r-s)^{\frac12 - H} r^{H-\frac12}\mathbb{E}\left[D_r^j G_j(T,t)\Big|\mathcal{F}_t\right] dr\right) ds\,.
\end{aligned}$$

Using Proposition 1.2.8 of [27], we have

$$\begin{aligned}
I_1 &= \int_0^t \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_s^t (r-s)^{\frac12 - H} r^{H-\frac12} D_r^j \mathbb{E}\left[G_j(T,t)\Big|\mathcal{F}_t\right] dr\right) ds \\
&= \int_0^t \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_s^t (r-s)^{\frac12 - H} r^{H-\frac12} D_r^j \mathbb{E}\left[P(t)\sigma_u^{j,*}(t)\Big|\mathcal{F}_t\right] dr\right) ds \\
&= \int_0^t \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_s^t (r-s)^{\frac12 - H} r^{H-\frac12} D_r^j \left(p(t)\sigma_u^{j,*}(t)\right) dr\right) ds\,.
\end{aligned} \tag{5.16}$$

I_2 is computed as follows

$$\begin{aligned}
I_2 &= \int_0^T \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_{t\vee s}^T (r-s)^{\frac12 - H} r^{H-\frac12}\mathbb{E}\left[D_r^j \left(P(t)\sigma_u^{j,*}(t)\right)\Big|\mathcal{F}_t\right] dr\right) ds \\
&= \sigma_u^{j,*}(t)\int_0^T \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_{t\vee s}^T (r-s)^{\frac12 - H} r^{H-\frac12}\mathbb{E}\left[D_r^j P(t)\Big|\mathcal{F}_t\right] dr\right) ds\,.
\end{aligned} \tag{5.17}$$

Now let us discuss $p(t)$. First we have

$$\begin{aligned}
p(t) &= \left(\Phi^\top\right)^{-1}(t)\mathbb{E}^{\mathcal{F}_t}\left[\int_0^T \Phi^\top(s) l_x^*(s) ds + \Phi^\top(T) h_x(x^*(T))\right] \\
&\quad -\left(\Phi^\top\right)^{-1}(t)\int_0^t \Phi^\top(s) l_x^*(s) ds\,.
\end{aligned}$$

Since $\left(\mathbb{E}^{\mathcal{F}_t}\left[\int_0^T \Phi^\top(s)l_x^*(s)ds + \Phi^\top(T)h_x(x^*(T))\right], 0\le t\le T\right)$ is a square integrable martingale with respect to the filtration $\mathcal{F}_t$ generated by the standard Brownian motion W, we have

$$\mathbb{E}^{\mathcal{F}_t}\left[\int_0^T \Phi^\top(s)l_x^*(s)ds + \Phi^\top(T)h_x(x^*(T))\right] = -\sum_{j=1}^m \int_0^t \Phi^\top(s)q_j(s)dW_j(s)\,,$$

where the introduction of the factor $-\Phi^\top(s)$ is to simplify the equation obtained subsequently. Therefore,

$$p(t) \quad = \quad -\left(\Phi^\top\right)^{-1}(t)\sum_{j=1}^m \int_0^t \Phi^\top(s)q_j(s)dW_j(s) - \left(\Phi^\top\right)^{-1}(t)\int_0^t \Phi^\top(s)l_x^*(s)ds\,.$$

Denote

$$K_t = -\sum_{j=1}^m \int_0^t \Phi^\top(s)q_j(s)dW_j(s)$$

and

$$A_t = \int_0^t \Phi^\top(s)l_x^*(s)ds\,.$$

Then by Proposition 2.7 and equation (4.44), we have

$$\begin{aligned} dp(t) &= d\Phi^{\top,-1}(t)\cdot K_t + \Phi^{\top,-1}(t)\cdot dK_t - d\Phi^{\top,-1}(t)\cdot A(t) - \Phi^{\top,-1}(t)\cdot dA(t) \\ &= -b_x^{*\top}(t)p(t)dt - l_x^*(t)dt - \sigma_x^{*\top}(t)p(t)d^\circ B^H(t) - q(t)dW(t), \end{aligned}$$

It is clear that

$$p(T) = h_x(x^*(T)),$$

Therefore, $p(t)$ satisfies the following backward stochastic differential equation

$$\begin{cases} dp(t) = -b_x^{*\top}(t)p(t)dt - l_x^*(t)dt - \sigma_x^{*\top}(t)p(t)d^\circ B^H(t) + \sum_{j=1}^m q_j(t)dW_j(t)\,, & 0\le t\le T \\ p(T) = h_x(x^*(T)) \end{cases} \tag{5.18}$$

Or

$$p(T) = h_x(x^*(T)) + \int_t^T \left[b_x^{*\top}(s)p(s) - l_x^*(s)\right]ds + \int_t^T \sigma_x^{*\top}(s)p(s)d^\circ B^H(s) + \int_t^T q_j(s)dW_j(s)\,.$$

Combining (5.14), (5.15), (5.16), and (5.17) we have

Theorem 5.7 *Let the assumptions (H4) and (H5) be satisfied. If (u^*, x^*) is an optimal pair of stochastic control problems (5.2)-(5.4). Then (u^*, x^*) satisfies*

the following system of equations.

$$
\begin{cases}
dx(t) = b(t, x(t), u(t))dt + \sum\limits_{j=1}^{m} \sigma^j(t, x(t), u(t))d^\circ B_j^H(t)\,, \\
x(0) = x_0\,; \\
\\
dp(t) = -b_x^{*\top}(t)p(t)dt - l_x^*(t)dt - \sigma_x^{*\top}(t)p(t)d^\circ B^H(t) + \sum_{j=1}^m q_j(t)dW_j(t)\,, \\
p(T) = h_x(x^*(T))\,; \\
\\
\sum_{j=1}^m \sigma_u^{j,*\top}(t)p(t) = 0\,; \\
\\
b_u^{*\top}(t)p(t) + l_u^*(t) + \int_0^t \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_s^t (r-s)^{\frac{1}{2}-H} r^{H-\frac{1}{2}} D_r^j\left(p(t)\sigma_r^{j,*}(t)\right)dr\right)ds \\
\quad +\sigma_u^{j,*}(t)\int_0^T \phi_{1,H}(t-s)\left(\frac{d}{ds}\int_{t\vee s}^T (r-s)^{\frac{1}{2}-H} r^{H-\frac{1}{2}}\mathbb{E}\left[D_r^j P(t)\middle|\mathcal{F}_t\right]dr\right)ds = 0\,,
\end{cases}
\tag{5.19}
$$

where

$$
P(t) = \left(\Phi^\top\right)^{-1}(t)\left[\int_t^T \Phi^\top(s)l_x^*(s)ds + \Phi^\top(T)h_x(x^*(T))\right].
$$

References

[1] Biagini, F., Hu, Y., Øksendal, B. and Zhang, T. *Stochastic calculus for fractional Brownian motion and applications.* Springer, 2008.

[2] Biagini, F., Hu, Y., Øksendal, B. and Sulem, A. *A stochastic maximum principle for processes driven by fractional Brownian motion.* Stochastic Process. Appl. 100 (2002), 233–253.

[3] Benes, V. E.; Shepp, L. A. and Witsenhausen, H. S. *Some solvable stochastic control problems.* Stochastics 4 (1980/81), no. 1, 39-83.

[4] Bismut, J. M. *Conjugate convex functions in optimal stochastic control.* J. Math. Anal. Appl., 44 (1973), 384-404.

[5] Bismut, J. M. *An introductory approach to duality in optimal stochastic control.* SIAM Rev., 20 (1978), 62-78.

[6] Bismut, J. M. *Mécanique aléatoire.* Lecture Notes in Mathematics, 866. Springer, 1981.

[7] Coutin, L. and Qian, Z. *Stochastic analysis, rough path analysis and fractional Brownian motions.* Probab. Theory Related Fields **122** (2002), 108–140.

[8] Cannarsa, P. and Sinestrari, C. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control.* Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser, 2004.

[9] Dellacherie, C. and Meyer, P.-A. Probabilities and potential. North-Holland Publishing Co., Amsterdam-New York; North-Holland Publishing Co., Amsterdam-New York, 1978.

[10] Duncan, T., Hu, Y. and Pasik-Duncan, B. *Stochastic calculus for fractional Brownian motion I. Theory*. SIAM J. Control Optim., 38(2) (2000), 582-612.

[11] Fleming, W. H. and Soner, H. M. *Controlled Markov processes and viscosity solutions*. Second edition. Stochastic Modelling and Applied Probability, 25. Springer, 2006.

[12] Friz, P. K. and Victoir, N. B. *Multidimensional stochastic processes as rough paths. Theory and applications*. Cambridge Studies in Advanced Mathematics, 120. Cambridge University Press, Cambridge, 2010.

[13] Haussmann, U. G. *General necessary conditions for optimal control of stochastic system*. Math. Programm. Stud., 6 (1976), 34–48.

[14] Hu, Y. *Integral transformations and anticipative calculus for fractional Brownian motions*. Memoirs of the American mathematical society, Number 825, 2005.

[15] Hu, Y. *Optimal consumption and portfolio in a market where the volatility is driven by fractional Brownian motion*. Probability, Finance and Insurance. Ed. Lai, T.L. et al. World Scientific Publishing. 164-173.

[16] Hu, Y. *Multiple integrals and expansion of solutions of differential equations driven by rough paths and by fractional Brownian motions*. To appear in Stochastics An International Journal of Probability and Stochastic Processes.

[17] Hu, Y. and Nualart, D. *Differential equation driven by Hölder continuous functions of order greater than* 1/2. in *The Abel Symposium on Stochastic Analysis*, 399-423. Springer, 2007.

[18] Hu, Y. and Nualart, D. *Rough path analysis via fractional calculus*. Trans. Amer. Math. Soc. 361 (2009), no. 5, 2689–2718.

[19] Hu, Y., Nualart, D. and Song, X. *Malliavin calculus for backward stochastic differential equations and application to numerical schemes*. The Annals of Applied Probability 21 (2011), no. 6, 2379-2423.

[20] Hu, Y. and Øksendal, B. *Partial information linear quadratic control for jump diffusions*. SIAM Journal of Control and Optimization, 47 (2008), 1744-1761.

[21] Hu, Y., Øksendal, B. and Sulem, A. *Optimal consumption and portfolio in a Black-Scholes market driven by fractional Brownian motion*. Infin. Dimens. Anal. Quantum Probab. Relat. Top. 6 (2003), no. 4, 519-536.

[22] Hu, Y. and Zhou, X. *Stochastic control for linear systems driven by fractional noises*. SIAM J. Control Optim. 43 (2005), 2245-2277.

[23] Kushner, H. J. *Necessary conditions for continuous parameter stochastic optimization problems*. SIAM J. Control, 10 (1972), 550-565.

[24] Lions, P. L. *Generalized solutions of Hamilton-Jacobi equations*. Research Notes in Mathematics, 69. Pitman Advanced Publishing Program, 1982.

[25] Lyons, T. and Qian, Z.M. *System Control and Rough Paths*. Clarendon Press. Oxford, 2002.

[26] Nisio, M. *Lectures on stochastic control theory*. ISI Lecture Notes, 9. Macmillan Co. of India, 1981.

[27] Nualart, D. *The Malliavin Calculus and Related Topics*. Springer, 2006.

[28] Nualart, D. and Rascanu, S. *Differential equations driven by fractional Brownian motion*. Collect. Math., 53 (2002), 55-81.

[29] Peng, S. *A general stochastic maximum principle for optimal control problems*. SIAM J. Control, 28 (1990), 966-979.

[30] Revuz, D. and Yor, M. *Continuous Martingales and Brownian Motion*. Springer, 1994

[31] Samko S. G., Kilbas A. A. and Marichev O. I. *Fractional Integrals and Derivatives. Theory and Applications*. Gordon and Breach, 1993.

[32] Wonham, W. M. *On the separation theorem of stochastic control*. SIAM J. Control, 6 (1968), 312-326.

[33] Young, L. C. *An inequality of the Hölder type connected with Stieltjes integration*. Acta Math. 67 (1936) 251-282.

[34] Zähle, M. *Integration with respect to fractal functions and stochastic calculus. I* . Prob. Theory Relat. Fields 111 (1998) 333-374.

[35] Yong, J. and Zhou, X. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer, 1999.